

14 March 2008 | S10

Science





COVER

Tomato (*Solanum lycopersicum*) domestication and selection resulted in an explosion of different fruit shapes and sizes, as exemplified by heirloom varieties. Wild ancestors of tomatoes such as *S. pimpinellifolium*, the smallest fruit shown here, were round (middle row, second from left). See page 1527.

Image: Kelly Krause/Science;
photos: Esther van der Knaap

DEPARTMENTS

- 1451 Science Online
- 1453 This Week in Science
- 1459 Editors' Choice
- 1464 Contact Science
- 1465 Random Samples
- 1467 Newsmakers
- 1547 New Products
- 1548 Science Careers

EDITORIAL

- 1457 The IPCC Must Maintain Rigor
by Susan Solomon and Martin Manning

NEWS OF THE WEEK

- Two Papers From Korean Lab Found to Lack 'Scientific Truth' 1468
- \$300 Million in Private Money for New Investigators 1469
- Physicist Wins Open Illinois Seat 1470
- Indian Government Offers Helping Hand to Women Scientists 1470
- Hiding From Biting Insects in Plain Scent 1471
>> Science Express Report by M. Ditzén et al.

SCIENCESCOPE

- Proposed Frog Ban Makes a Splash 1472
- Spanish Scientists Question Accreditation Plan 1472
- U.K.'s Royal Society Ventures Into Funding Start-Up Companies 1473

NEWS FOCUS

- Dueling Visions for a Hungry World 1474
- NSF Fellowships Called Powerful Tool for Building the Pipeline 1477
- Blood-Matching Goes Genetic 1478
- Cryptologists Cook Up Some Hash for New 'Bake-Off' Hash of the Future? 1480



1474

LETTERS

- The Right Pitch for Saxophonists P. King 1483
- Second Basket's Negative Impact C.-C. Carbon
- Going Public with the Scientific Process R. Cronje
- Hoyle's Role in B²FH G. Burbridge

- CORRECTIONS AND CLARIFICATIONS 1484

BOOKS ET AL.

- Reading Popular Physics Disciplinary Skirmishes and Textual Strategies E. Leane, reviewed by D. Kaiser 1485
- Elizabeth Blackburn and the Story of Telomeres Deciphering the Ends of DNA C. Brady, reviewed by M. A. Goldman 1486

POLICY FORUM

- Inappropriate Use and Portrayal of Chimpanzees S. R. Ross et al. 1487

PERSPECTIVES

- Observing Our Origins F. J. Ciesla 1488
>> Report p. 1504
- Customized Signaling Circuits P. M. Pryciak 1489
>> Report p. 1539
- The New Diamond Age? P. W. May 1490
- This Coincidence Cannot Be Accidental D. J. Scalapino 1492
>> Report p. 1509
- Standing on the Shoulders of Giants F. Coward 1493
- A Quiescent Working Memory S. Fusi 1495
>> Report p. 1543



1485

CONTENTS continued >>



SCIENCE EXPRESS

www.sciencexpress.org

PHYSICS

Coherent Dynamics of a Single Spin Interacting with an Adjustable Spin Bath

R. Hanson, V. V. Dobrovitski, A. E. Feiguin, O. Gywat, D. D. Awschalom

Simulations successfully show how the spin of a nitrogen vacancy in diamond is coupled to those of surrounding nitrogen impurities and how coherence between them is lost.

[10.1126/science.1155400](https://doi.org/10.1126/science.1155400)

CHEMISTRY

Gate-Variable Optical Transitions in Graphene

F. Wang et al.

Application of electrical biases to single or double layers of graphene change its infrared reflectivity, mimicking aspects of transistors and opening up optoelectronic applications.

[10.1126/science.1152793](https://doi.org/10.1126/science.1152793)

CLIMATE CHANGE

Impact of Artificial Reservoir Water Impoundment on Global Sea Level

B. F. Chao, Y. H. Wu, Y. S. Li

Accounting for water impounded globally in artificial lakes that were filled during the past 80 years raises estimates of natural contributions to recent sea level.

[10.1126/science.1154580](https://doi.org/10.1126/science.1154580)

MOLECULAR BIOLOGY

Nutritional Control of Reproductive Status in Honeybees via DNA Methylation

R. Kucharski, J. Maleszka, S. Foret, R. Maleszka

Epigenetic modifications that involve methylation cause female honeybee larvae to become queens rather than workers when they are fed royal jelly.

[10.1126/science.1153069](https://doi.org/10.1126/science.1153069)

NEUROSCIENCE

Insect Odorant Receptors Are Molecular Targets of the Insect Repellent DEET

M. Ditzgen, M. Pellegrino, L. B. Vosshall

The widely used insect repellent DEET acts by inhibiting olfactory neurons that respond to odors such as those that attract insects to their hosts.

>> [News story p. 1471](#)

[10.1126/science.1153121](https://doi.org/10.1126/science.1153121)

REVIEW

ANTHROPOLOGY

The Late Pleistocene Dispersal of Modern Humans in the Americas 1497

T. Goebel, M. R. Waters, D. H. O'Rourke

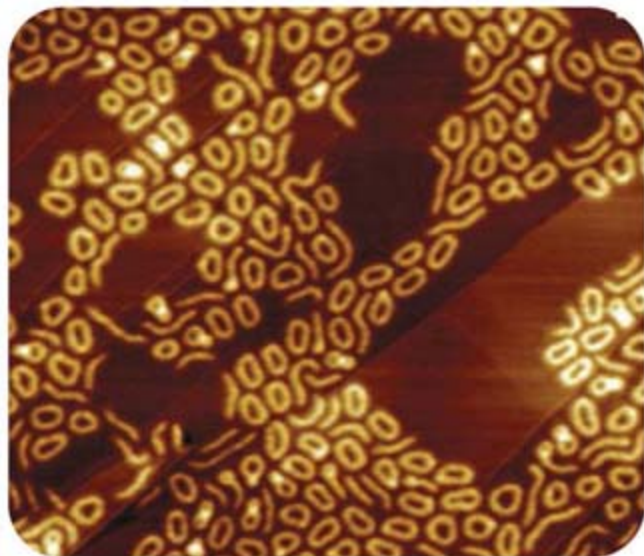
BREVIA

ECOLOGY

Predators Induce Cloning in Echinoderm Larvae 1503

D. Vaughn and R. R. Strathmann

Sand dollar larvae respond to mucus from fish predators by rapid asexual reproduction, producing an increased number of smaller individuals that may be less visible.



[1512](https://doi.org/10.1126/science.1151212)

REPORTS

ASTRONOMY

Organic Molecules and Water in the Planet Formation Region of Young Circumstellar Disks 1504

J. S. Carr and J. R. Najita

Simple organic molecules and water are abundant in the inner disk of a star like our early Sun, implying that organic synthesis is occurring there.

>> [Perspective p. 1488](#)

PHYSICS

Superconductivity in Hydrogen Dominant Materials: Silane 1506

M. I. Eremets, I. A. Trojan, S. A. Medvedev, J. S. Tse, Y. Yao

Under pressure, the insulator silane (SiH₄) transforms to a metallic phase and, at even higher pressures and low temperatures, becomes superconducting.

PHYSICS

Energy Gaps and Kohn Anomalies in Elemental Superconductors 1509

P. Aynajian et al.

High-resolution neutron-scattering experiments reveal behavior in pure lead and niobium superconductors beyond that described by the standard theoretical framework.

>> [Perspective p. 1492](#)

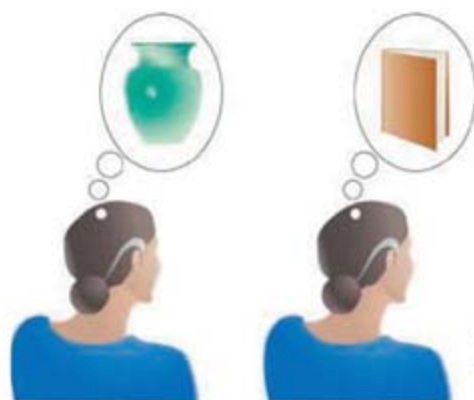
CHEMISTRY

Synthesis of Macrocyclic Copolymer Brushes and Their Self-Assembly into Supramolecular Tubes 1512

M. Schappacher and A. Deffieux

A three-block polymer can generate long cyclic polymers in solution, which can be further functionalized to form polymer brushes and tubular assemblies.

CONTENTS [continued >>](#)



1495 &
1543

REPORTS CONTINUED...

GEOCHEMISTRY

Splitting of the 520-Kilometer Seismic Discontinuity and Chemical Heterogeneity in the Mantle 1515

A. Saikia, D. J. Frost, D. C. Rubie

A reaction in calcium-rich rocks in Earth's mantle can explain a seismic signal that varies geographically and with depth, providing a means to map mantle compositions.

GEOCHEMISTRY

The Chlorine Isotope Composition of Earth's Mantle 1518

M. Bonifacie et al.

Analysis of ocean basalts imply that Earth's mantle has a different chlorine isotopic ratio than crust or seawater, perhaps reflecting a late addition of material to Earth.

PALEONTOLOGY

Caribbean Reef Development Was Independent of Coral Diversity over 28 Million Years 1521

K. G. Johnson, J. B. C. Jackson, A. F. Budd

An analysis of fossil and modern Caribbean corals shows that, for the last 28 million years, coral reef growth and persistence have not required high coral diversity.

BIOCHEMISTRY

Amyloid Fibrils of the HET-s(218–289) Prion Form a β Solenoid with a Triangular Hydrophobic Core 1523

C. Wasmer et al.

A structural model of a yeast prion shows that the amyloid fibrils form a left-handed β solenoid stabilized by hydrophobic and polar interactions and salt bridges.

GENETICS

A Retrotransposon-Mediated Gene Duplication Underlies Morphological Variation of Tomato Fruit 1527

H. Xiao et al.

The gene causing elongated tomatoes arose from an unusual, transposon-mediated duplication that provided a new regulatory environment that increased its expression.

GENETICS

Evidence for Karyogamy and Exchange of Genetic Material in the Binucleate Intestinal Parasite 1530

Giardia intestinalis

M. K. Poxleitner et al.

The unusual lack of accumulated mutations in asexual *Giardia* can be explained by the exchange of plasmid DNA between its two nuclei during the cyst phase.

MICROBIOLOGY

Direct Visualization of Horizontal Gene Transfer 1533

A. Babić et al.

Visualization of DNA exchange between two bacteria reveals that the process is highly efficient, is mediated by the pilus, and occurs about once per replication cycle.

MEDICINE

Neurokinin 1 Receptor Antagonism as a Possible Therapy for Alcoholism 1536

D. T. George et al.

A drug that inhibits a neural signaling pathway linked to behavioral stress may be a useful therapy in preventing relapse in alcoholics.

CELL BIOLOGY

Using Engineered Scaffold Interactions to Reshape MAP Kinase Pathway Signaling Dynamics 1539

C. J. Bashor, N. C. Helman, S. Yan, W. A. Lim

A yeast signaling pathway acquires new regulatory properties (such as adaptation) when additional protein-protein interaction sites are engineered into a scaffolding protein.

>> *Perspective p. 1489*

NEUROSCIENCE

Synaptic Theory of Working Memory 1543

G. Mongillo, O. Barak, M. Tsodyks

Stronger synapses induced by calcium currents are responsible for working memory rather than the more metabolically expensive action potential firing, as had been thought. >> *Perspective p. 1495*

SPECIAL FEATURE

Financing Your Research From Industry

www.sciencereers.org

Playing Well With Industry	1548
Business Financing for Your Research	1549
Finding Industry Funding	1550



SCIENCE (ISSN 0036-8075) is published weekly on Friday, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals Mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. Copyright © 2008 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership and subscription (51 issues): \$144 (\$74 allocated to subscription). Domestic institutional subscription (51 issues): \$770; Foreign postage extra: Mexico, Caribbean (surface mail) \$55; other countries (air assist delivery) \$85. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request. GST #1254 88122. Publications Mail Agreement Number 1069624. SCIENCE is printed on 30 percent post-consumer recycled paper. Printed in the U.S.A.

Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. Postmaster: Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090-6178. Single-copy sales: \$10.00 current issue, \$15.00 back issue prepaid includes surface postage; bulk rates on request. Authorization to photocopy material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act is granted by AAAS to libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that \$20.00 per article is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923. The identification code for Science is 0036-8075. Science is indexed in the Reader's Guide to Periodical Literature and in several specialized indexes.

CONTENTS continued >>>



Water waster?

SCIENCE NOW

www.sciencenow.org DAILY NEWS COVERAGE

Future Cars May Save Gas But Waste Water
Plug-in vehicles could strain resources of drought-stricken regions.

He Heard, She Heard
Girls and boys use different parts of their brains to process language.

Mathematicians Debate the Hole Truth
Who was the first to improve on a widely used classic formula?



From drifter to researcher.

SCIENCE CAREERS

www.sciencereers.org CAREER RESOURCES FOR SCIENTISTS

From the Archives: The Drifter
I. Chen
After 2 years of wandering, Gary Ruvkun settled into a productive research career.

>> See also *Financing Your Research From Industry* feature, p. 1548



Find a drug, any drug.

SCIENCE SIGNALING

www.stke.org THE SIGNAL TRANSDUCTION KNOWLEDGE ENVIRONMENT

- ST NETWATCH: DrugBank**
DrugBank is a useful tool for researchers using pharmacological agents to study signaling pathways; in Bioinformatics Resources.
- ST NETWATCH: MINT, the Molecular INTERaction Database**
Explore your favorite protein's potential binding partners with the MINT database; in Protein Databases.
- ST NETWATCH: Ask a Scientist**
This Web site from the Howard Hughes Medical Institute lets users submit their questions directly to scientists; in Educator Sites.



SCIENCE PODCAST

Download the 14 March *Science* Podcast to hear about predator-induced sand dollar cloning, the composition of protoplanetary disks, a worldwide assessment of agricultural science, and more.
www.sciencemag.org/about/podcast.dtl

Separate individual or institutional subscriptions to these products may be required for full-text access.



<< Maintaining Duplicate Nuclei

The intestinal parasite *Giardia* has two nuclei in its vegetatively growing form (the trophozoite), and both nuclei seem to be fully functional and essentially indistinguishable. Remarkably, despite no previous evidence for genetic exchange between the two nuclei, they somehow appear to contain essentially identical copies of the genome. Poxleitner *et al.* (p. 1530) show that efficient DNA transfer occurs during encystations and provides a mechanism for maintaining homozygosity between the two nuclei.

Migration Dates to the New World

The Americas were the last major habitable continents to be settled by humans. Goebel *et al.* (p. 1497) review recent archaeological finds, more accurate dates, and genetic evidence that have changed some ideas about how this migration occurred. The data favor an initial migration along a coastal ice-free corridor about 15,000 years ago from populations that were in Beringia, including Alaska, and perhaps a second migration around 13,000 years ago that led to the Clovis culture in North America.

Protoplanetary Parts Lists

Our understanding of the initial conditions of our early solar system is based mainly on samples of material that has condensed and accumulated into grains, comets, asteroids, and planets. Carr and Najita (p. 1504; see the Perspective by Ciesla) present spectral observations obtained with Spitzer Space Telescope of the gas in the protoplanetary disk around the star AA Tauri, which is thought to be similar to our early Sun. Even the inner portion of this disk contains high concentrations of HCN, CO₂, C₂H₂, OH, and H₂O, which suggests that organic chemical reactions are ongoing. These results raise the possibility that even more complex molecules are synthesized in these regions, and may also help in understanding the origin of Earth's water budget.

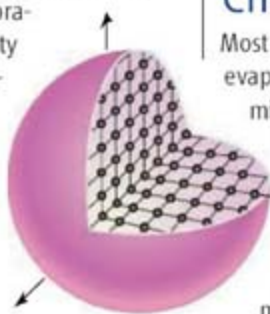
Hydrogen Surrogates Under Pressure

The issue of hydrogen metallization at extreme pressures is not only of fundamental interest but also bears on our understanding of the deep interior of gas giant planets. This issue is diffi-

cult to address experimentally in that pressure cells, which allow for detailed observations, cannot yet induce elemental hydrogen into its metallic phase. Eremets *et al.* (p. 1506) turned to compounds containing hydrogen—in this case, silanes—with the motivation that because the hydrogen is “chemically precompressed,” it should be easier to induce phase changes at lower pressure. Silane can indeed be pressured into a metallic phase and is also superconducting at lower temperature. The results suggest that hydrogen-rich compounds can be a good surrogate for studying elemental hydrogen.

Look Beyond BCS Theory?

The Bardeen-Cooper-Schrieffer (BCS) theory of low-temperature superconductivity in metals stands as a monument in solid-state physics, and its cornerstone is a pairing interaction between electrons mediated by lattice vibrations, or phonons. The availability of high-resolution neutron spectroscopy now allows measurement of phonon lifetimes from micro-electron volt resolution studies throughout the Brillouin zone. Aynajian *et al.* (p. 1509, published online 21 February; see the Perspective by Scalapino) studied the elemental superconductors lead and niobium with such methods and found results which are suggestive of many-body correlated effects that require explanations beyond the standard BCS framework.



Blocking Out Cyclic Polymers

Synthesis of cyclic polymers requires some way of bringing the ends of a linear polymer

together before competing reactions occur. For this reason, such reactions are usually performed under high dilution, and for long polymers, this requirement often leads to low yields. Schappacher and Deffieux (p. 1512) overcome a number of these limitations by synthesizing a triblock copolymer with a long middle block and two smaller end blocks with reactive groups needed for the cyclization reaction. The central block could also be decorated with side chains, which aided in visualization with atomic force microscopy. By using two monomers to decorate the cyclic polymers and by choosing a selective solvent for one of them, it was possible to form supramolecular tubular assemblies in solution.

Constraining Earth's Chlorine Budget

Most of Earth's chlorine is now in the crust (in evaporates and brines) and the oceans. Determining how and when it became concentrated there requires some assessment of the mantle chlorine concentration and isotopic value. Such measurements are often difficult—for example, volcanic rocks from the mantle, such as mid-ocean ridge basalts (MORBs), are rapidly contaminated by seawater. Bonifacie *et al.* (p. 1518) examined Cl stable isotopes in 22 MORB samples and found, in contrast to previous studies, that the mantle chlorine has relatively less ³⁷Cl than ³⁶Cl. This difference implies either that fractionation accompanied separation from the mantle, and later mixing in of altered ocean crust, or that the crust of the Earth contains isotopically different Cl from a late accretion event.

Continued on page 1455

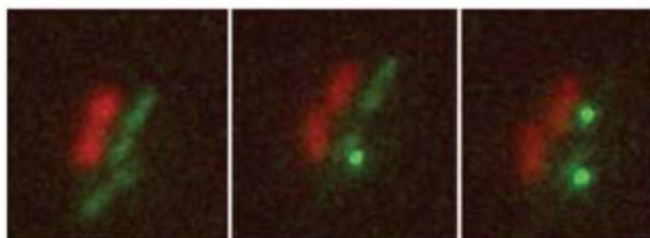
Continued from page 1453

Yeast Prion Protein Structure Revealed

Prion proteins are linked to several diseases, including bovine spongiform encephalopathy in cattle, scrapie in sheep, and Creutzfeldt-Jakob disease in humans. Infectious prion-like proteins that also form amyloid fibrils are found in yeast and other fungi. Based on solid-state nuclear magnetic resonance data, **Wasmer *et al.*** (p. 1523) provide a structural model of amyloid fibrils from the prion-forming domain of the fungal HET-s protein. An earlier preliminary model gave no information on intermolecular β -sheet propagation. The current model shows that the amyloid fibrils form a left-handed β -solenoid with two windings per molecule that is stabilized by hydrophobic and polar interactions and salt bridges.

Bacterial Conjugation in Living Color

Although conjugation in bacteria has been studied for decades, there are many outstanding issues about how it works. By using a direct visualization technique, **Babić *et al.*** (p. 1553) show that DNA is transferred through the pilus and efficiently incorporated (>95%) into the recipient's bacterial chromosome; unrecombined DNA is degraded. Tracking of the acquired DNA showed that the resulting chromosomes appeared to recombine frequently, about once per cell generation, and was able to generate a diverse set of progeny from a single ancestor.



Stress Relief in Alcoholism

Despite the success of psychosocial interventions such as Alcoholics Anonymous, chronic alcohol abuse remains a significant public health problem. Noting that stress is a relapse trigger in alcoholism, **George *et al.*** (p. 1536, published online 14 February) explored whether pharmacological inhibition of the neurokinin 1 receptor (NK1R), a mediator of the brain's stress response, would alleviate symptoms associated with alcohol dependence. Mice genetically deficient in NK1R consumed less alcohol than controls. A drug that antagonizes NK1R was used in a small, controlled study of recently detoxified alcoholic inpatients and showed promising activity, such as reducing alcohol cravings. However, these latter results need to be confirmed in much larger, longer-term studies.

Synthetic Biology in Action

One goal of synthetic biology is to remodel intracellular processes at will. **Bashor *et al.*** (p. 1539; see the Perspective by **Pryciak**) wanted to engineer new regulatory properties into the well-studied mitogen-activated protein kinase pathway in yeast that mediates response to mating pheromones. By creating proteins with customized protein interaction domains, which altered the association of signaling proteins with a key scaffold protein that assembles proteins that participate in the signaling pathway, the authors changed complex regulatory properties of the signaling pathway. A gradual dose response could be modified into a switchlike all-or-none response, the time course of system response could be varied, or adaptation (a process in which output diminishes even in the presence of continued stimulus) could be added.

A Working Model of Working Memory

Working memory, the system that enables us to hold information temporarily for processing purposes, is thought to be stored primarily in the form of elevated, persistent neuronal firing rates. **Mongillo *et al.*** (p. 1543; see the Perspective by **Fusi**) now suggest that working memory may instead be stored in the short-term facilitation caused by a temporary calcium increase in synapses from neurons that code for a particular memory. Short-term facilitation on the necessary time scale has recently been observed in prefrontal cortex, lending itself to memory maintenance on the scale of seconds. During the memory span, and also during recall, neuronal activity then shows population spikes. If memory is indeed stored in the form of synaptic facilitation, few spikes need to be expended in maintaining it, which would reduce the metabolic cost of working memory.

CREDIT: BABIĆ ET AL.

Warming Island, GREENLAND Expedition September 16-27, 2008



Join explorer **Dennis Schmitt** as he returns to East Greenland and his discovery—a three-finger-shaped island in East Greenland now named **Warming Island**—a compelling indicator of the rapid speed of global warming.

In **Reykjavik, Iceland**, we will board the 50-passenger expedition vessel, M/V **Aleksey Maryshev**, and go north across the Denmark Strait and above the Arctic Circle to the coast of East Greenland.



Blue whales feed in the rich waters, and orcas (killer whales), white-beaked dolphins, and sea birds may be seen.

We will visit **Scoresby Sund**, the longest fjord in the world, and at **Cape Hofmann Halvo** we will look for musk oxen. Remains of remote Inuit villages will be of interest, as will seals and other wildlife—all against the stunning glaciers and peaks of coastal Greenland. From \$5,745 + air.



For a detailed brochure,
please call (800) 252-4910

AAAS Travels

17050 Montebello Road
Cupertino, California 95014

Email: AAASInfo@betchartexpeditions.com



Susan Solomon is co-chair of IPCC Working Group I and a senior scientist at the Chemical Sciences Division, Earth System Research Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO.



Martin Manning is the former head of the IPCC Working Group I Technical Support Unit and a professor at Victoria University, Wellington, New Zealand.

The IPCC Must Maintain Its Rigor

LAST YEAR, THE INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC) WAS awarded the Nobel Peace Prize along with Albert Gore Jr., sending a strong message about the importance of the world's future climate. Indeed, for two decades, international scientists and policy-makers contributing to the IPCC process have provided assessments of climate change science, impacts, and mitigation, addressing one of the most far-reaching and complex challenges that society has ever faced. Yet this is no time for the IPCC to rest on its laurels. The climate system continues to change and science continues to improve, so policy must be kept current with our best understanding. Reformulating the science/policy interface should be considered and be open to change but must acknowledge lessons from the past. The factors that have been critical to the success of the IPCC need to be preserved if a rigorous scientific basis is to continue to inform the growing challenge of decision-making on climate change.

The IPCC's assessments have been highly credible to both policy-makers and scientists largely because, despite differences in language and approach between these communities, both groups recognize the importance of multiple reviews of drafts in an open process in which diverse authors, reviewers, and governments do not hide behind a cloak of anonymity. The successive reviews and revisions of complex material, with the broad inputs of many, take time. For instance, after scientists and governments defined its scope, the last Working Group I assessment took 3 years to be developed by 152 authors and then reviewed by more than 600 experts along with dozens of governments. Any move toward more rapid products risks incomplete identification of the range of justifiable views and a consequent reduction of the rigor, clarity, and robustness of the consensus.

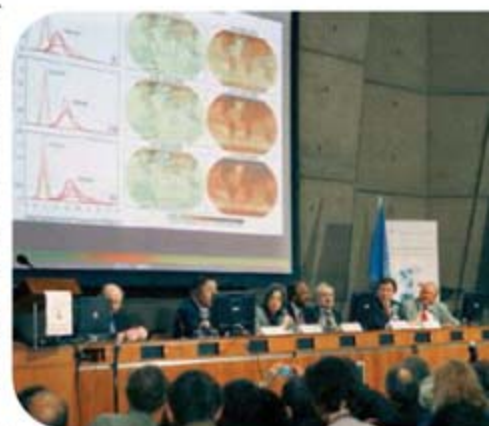
Moreover, completion of an IPCC assessment report requires a demanding line-by-line approval of its summary that is critical for its value to policy-makers. This process ensures that key conclusions are accepted by all governments and expressed in language that is both scientifically accurate and useful to policy. It requires Working Group co-chairs and authors with the stature and expertise needed to justify their findings and levels of scientific confidence. The utility of the IPCC also depends on its direct relevance to climate policy decisions, and this sharp clarity of purpose requires that the IPCC avoid becoming entrained in many aspects of broader global change and sustainable development issues.

IPCC meetings of experts play roles in addressing specific issues such as drought or climate sensitivity, but they do not include the crucial and interactive aspects of review and approval and hence cannot serve either the science or policy communities in the same way as assessment reports. IPCC special reports follow the same process as a comprehensive assessment and can be highly valuable, but cover limited subtopics (such as aviation). Thus, only periodic comprehensive IPCC assessments can span the breadth of policy-relevant issues and research advances, and they must be retained for a proper balance between science push and policy pull.

The IPCC does not plan or carry out research, and this separation between research and assessment is essential if the IPCC is to be an objective assessor. The mandate of the IPCC is to evaluate information that must be independently documented, primarily as peer-reviewed literature. The planning and coordination of international research are best carried out by organizations such as the World Climate Research Programme, the International Geosphere-Biosphere Programme, and the International Human Dimensions Programme. These bodies often consider IPCC assessments and help provide the means for the scientific community to produce related research.

In physical climate science, there are pressing needs for further information on coupled climate system changes. There will very likely soon be further advances in understanding changes in temperature, rainfall, hurricanes and other extreme events, sea-level rise, regional modeling, the carbon cycle, and other climate-forcing factors. These and other developments are certain to provide rich grounds for a new comprehensive physical science assessment within about 5 to 7 years that can, if conducted using an appropriately rigorous process, be the foundation for the most appropriate next steps in global climate policy.

— Susan Solomon and Martin Manning





Skull embedded in flowstone.

ANTHROPOLOGY

Early People of Palau

Diminutive fossils of *Homo sapiens*, perhaps representing several tens of individuals, have been found in two caves in Palau. The fossils, described by Berger *et al.*, range in age from about 3000 to 1500 years; humans are thought to have arrived on the island from the Philippines (700 km to the west) about 1000 years earlier. The fossils include several complete, small crania still encased in flowstone. Preliminary measurements suggest a brain size near the low end of extant *Homo sapiens* and close to that of *Homo erectus*. Measurements of multiple postcranial bones imply a corresponding body size comparable to those of the smallest known *H. sapiens* and the early hominin Lucy. Although many traits are consistent with *H. sapiens*, some fossils also exhibit more primitive traits, including minimal chins and an enlarged brow ridge. These traits and some aspects of the teeth and the small body size are similar to those seen in the older, enigmatic diminutive fossils found recently on the nearby island of Flores, which in turn have been ascribed to a relict population of an earlier *Homo* species. Although any relation between these fossils is not clear, the sample on Palau is further evidence of the extremes in size and characteristics that may develop in isolated island human populations. — BH

PLoS One 10.1371/journal.pone.0001780 (2008).

CHEMISTRY

Cool Clustering

The decomposition of a precursor within a polymeric medium can yield metal nanoclusters with narrow size distribution and uniform morphology. Gazit *et al.* explored the impact of this approach, performing the synthesis below the glass transition temperature (T_g) of the polymer, thereby precluding large-scale motions of the polymer segments. For iron nanoparticles within a polystyrene (PS) matrix, a large reduction in temperature shifted size and morphology from 70-nm spherical particles to 50-by-3-nm needles, whereas intermediate temperatures produced 8-nm spheres. Formation rates scaled exponentially with temperature above and below T_g . In the case of a poly(methyl methacrylate) (PMMA) matrix, similar size and morphology changes occurred, but the rate of cluster growth below T_g proved insensitive to temperature. The authors attribute the discrepancy to the relatively strong coordinative bonding between cluster and polymer present in the PMMA matrix but absent from PS. Taken together, the results support tuning of polymer segment mobility as

an additional control variable in the synthesis of metal clusters. — MSL

Macromolecules 41, 10.1021/ma071816o (2008).

PLANT BIOLOGY

Leaf Production Plant

Leaf epidermis, usually a single layer of cells, negotiates the balance between environmental factors and internal physiology. Leaf epidermal cells synthesize a wax that serves as a protective surface barrier and also synthesize a variety of complex metabolites important in plant defense. Leaves of the Madagascar periwinkle (*Catharanthus roseus*), long known as an herbal medicine, synthesize the alkaloids vinblastine and vincristine, which have valuable anticancer activities. In the search for other useful alkaloids or metabolites, Murata *et al.* have exploited RNA analysis of the leaf epidermis. *C. roseus* epidermis is rich in a variety of biosynthetic pathways, including those for flavonoids, lipids, and pentacyclic triterpenes. Analysis of RNAs expressed in



Catharanthus roseus.

leaf epidermis revealed much more complexity of the biochemistry occurring in leaves than did similar analytic approaches using broader tissue samples of the plant. The diversity of biosynthetic pathways represented in the leaf epidermis, combined with the variable destinations of the products, makes this single layer of cells look like the busiest of production and distribution centers. — PJH

Plant Cell 10.1105/tpc.107.056630 (2008).

NEUROSCIENCE

More to Noradrenaline

Noradrenaline is a well-established neuromodulator in many parts of the brain. Its effects have usually been described in terms of its impact on cells or synapses. Using paired recordings from connected GABAergic interneurons in the hippocampus, Zsiros and Maccaferri investigated a novel type of noradrenergic modulation. Activation of β -adrenergic receptors decreased gap junction-dependent electrical coupling between inhibitory interneurons. A series of pharmacological interventions established that this effect was due to activation of the cAMP/protein kinase A signaling cascade. Electrical coupling is important for the coordination of interneuron activity, which may lead to synchronous firing and oscil-

Continued on page 1461

Continued from page 1459

lations of larger neuronal networks. Noradrenergic modulation of electrical coupling between connected inhibitory interneurons can thus have important effects on information transmission in the whole GABAergic inhibitory hippocampal network and hence on signal propagation throughout the hippocampus and the many other brain regions connected to it. — PRS

J. Neurosci. **28**, 1804 (2008).

ECOLOGY

Shared Prosperity

The relationships between biodiversity and ecosystem function (such as productivity and nutrient dynamics) have mostly been investigated with tractable ecological communities such as herbaceous vegetation. Now Potvin and Gotelli have extended such studies to simple tree communities, with an experiment on the effects of tree species



Forest plantation at 5 years.

diversity on yield, measured by growth in tree basal area. In a forest plantation in Panama, plots in which several species of tree seedlings were planted yielded 30 to 58% more growth than monocultures after 5 years. The increased yield resulted from increased growth in the mixed-species plots rather than

from mortality in the monocultures. The authors speculate that competition for light is greater in monocultures, implying that more effective partitioning of resources permits more biomass accumulation in the mixed-species plots. — AMS

Ecol. Lett. **11**, 217 (2008).

CHEMISTRY

Crystals of an Iron Nitride

Both industrial and enzymatic nitrogen reduction catalysts rely on iron centers. However, high valent molecular iron nitride complexes (Fe=N) have stubbornly eluded crystallographic characterization, in contrast to analogous terminal oxo structures (Fe=O). Vogel *et al.* now find that a tridentate ligand comprising three coordinating N-heterocyclic carbene moieties offers the solution. Reaction of the ligand (bearing either xylyl or mesityl groups for steric protection) with ferrous chloride followed by reduction with sodium amalgam and treatment with trimethylsilylazide yields a cationic Fe-N₃ complex that loses dinitrogen under xenon lamp photolysis to afford the terminal iron nitride. Air-stable purple crystals of the compound were characterized by x-ray diffraction, revealing a short FeN bond length of 1.53 Å. Mössbauer spectroscopy further suggested an iron center more electron-rich than a previously prepared phosphorus-coordinated iron nitride characterized in solution; the authors attribute the difference to π-donation from the carbene ligands. — JSY

Angew. Chem. Int. Ed. **47**, 10.1002/anie.200800600 (2008).

Science Signaling



<< Pulling on the TCR

The exact mechanism by which peptide-loaded major histocompatibility complex (pMHC) activates the T cell receptor (TCR) has been controversial. Effective TCR activation requires presentation by antigen-presenting cells—soluble pMHC cannot activate T cells. In the presence of endogenous pMHC, however, very low concentrations of agonist pMHC are required to activate TCR signaling, as are costimulatory interactions. Ma *et al.* anchored pMHCs to lipid bilayers or plastic surfaces and found that in the absence of nonagonist pMHC, a minimum of 1 to 10 monomeric agonist pMHCs (MCC-loaded pMHC) per T cell was sufficient to stimulate TCR signaling in T cells interacting with the artificial substrates. Furthermore, the T cell response (calcium signal) was not enhanced by the addition of nonagonist pMHC with the agonist pMHC or under conditions where the nonagonist pMHC and agonist pMHC were close enough to function as a dimer. When murine B cell lymphoma cells were loaded with various endogenous peptides and the MCC peptide, the production of interleukin-2 by the T cells was the same whether a costimulatory endogenous pMHC was present or not. In further experiments, adhesion was shown to be necessary for TCR triggering, as was the actin cytoskeleton. The authors propose a model whereby the agonist pMHC-TCR interaction leads to receptor deformation and activation when the cytoskeleton provides a pulling force on the complex. Endogenous pMHC-TCR interactions would be insufficiently strong to provide the necessary activation signal, and the complex would dissociate when “pulled on” by the cytoskeleton. — NRG

PLoS Biol. **6**, e43 (2008).

World's largest
c57BL/6 gene trap library
CLICK ON IT



UT
TEXAS INSTITUTE FOR GENOMIC MEDICINE

713-677-7429 | 888-377-TIGM

Galileo in the Vatican

Better late than never. Nearly 4 centuries after Galileo was persecuted for heresy, the Vatican will erect a statue of the pioneer astronomer in its garden.

Armed with only a Dutch "spyglass," Galileo shook the foundations of the prevailing view of the universe with discoveries about the moon and Jupiter as well as the heretical notion that Earth revolves around the sun. He stood trial in 1633, was forced to recant his discoveries, and remained under house arrest until his death in 1642.

It was not until 1979 that Pope John Paul II encouraged the clergy to reconsider the episode. Now, with the statue, "the Church wants to close the Galileo affair and reach a definitive understanding not only of his great legacy but also of the relationship between science and faith," Nicola Cabibbo, a nuclear physicist and head of the Pontifical Academy of Sciences, told *The Times* of London.



1870 painting of Galileo with secretary.

"I think that the statue is a way for the Church to get away [from] the Galileo affair without embarrassment," says astrophysicist Simone Recchi of the University of Trieste in Italy. But Recchi says it's not enough: To win the support of scientists, it must "concentrate on present problems—such as stem cells, contraceptives, euthanasia, abortion—and open a fair debate" about them.

Doctor Is Out

A German law designed to prevent people from impersonating physicians and police officers has gotten some U.S.-educated scientists in hot water. A half-dozen Max Planck Institute researchers have been charged with unauthorized use of the title "Dr."

Ian Baldwin, a director at the Max Planck Institute for Chemical Ecology in Jena, says he received a letter in January charging him with *Titelmissbrauch*. "I spent the weekend wondering whom I had killed," he says. Two other directors of the institute have also been charged, as have researchers in Dresden and Potsdam.

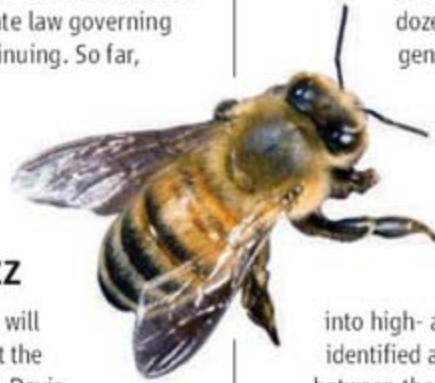
In Germany, only those who receive their doctorates in the E.U. are allowed to use Dr.,

which is considered a legal part of a name. Maximum penalty is a year in jail. On 6 March, authorities announced that holders of U.S. Ph.D.s will be allowed to use the title. But because of a loophole in state law governing Jena, Baldwin's case is continuing. So far, he's keeping his sense of humor. "If they have e-mail in jail, this could be productive," he says.

Homeward Buzz

This month, a famous family will return to its original home at the University of California (UC), Davis, after a stint at Arizona State University (ASU). A colony of 750,000 honey bees, now in its 32nd generation, was recently trucked to California.

Honey bee geneticist Robert E. Page Jr., who developed the colony 18 years ago at UC Davis, took the bees with him when he moved to ASU as founding director of its School of Life Sciences. But they were stressed by the variable Arizona climate, so now they're pollinat-



ing almonds en route to settling at Davis.

The country's oldest and most carefully characterized bee research colony, according to Page, has led to dozens of discoveries about bee genes, behavior, and development, all of which were made by selecting for one trait: pollen storage. Bees have an intricate social system based on how much pollen they store in their combs. By breeding bees into high- and low-storing stocks, "we've identified a whole suite of traits that differ between them," says Page. For example, high storers are better at associative learning.

The Davis colony has a new friend: the makers of Häagen-Dazs ice cream. "More than 40% of our flavors depend on honey bee pollination," says spokesperson Katty Pien. The company just gave UC Davis \$100,000 for research on pollination and on colony collapse disorder, which has caused agricultural havoc (*Science*, 18 May 2007, p. 970). Some of the money will also support a Häagen-Dazs Postdoctoral Fellow.

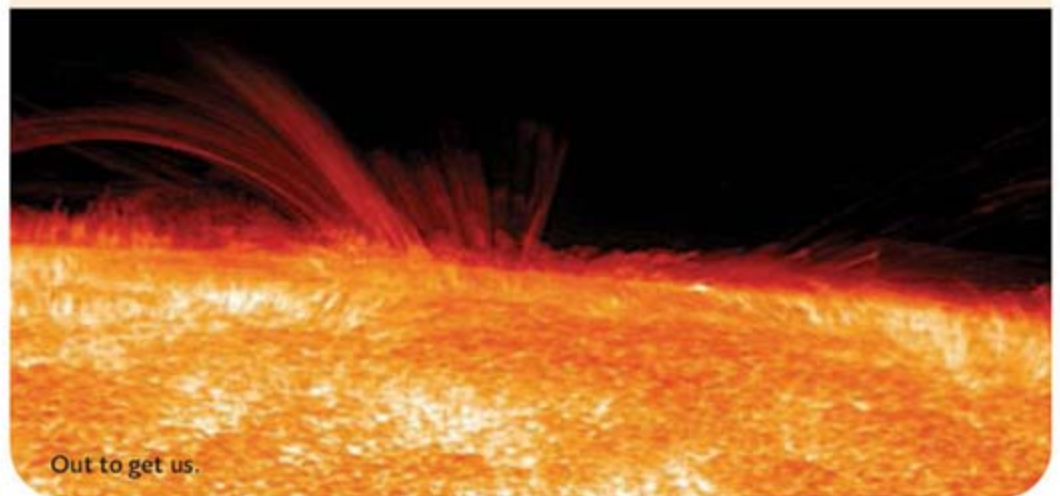
Out of the Frying Pan

Only 7.6 billion years from now, it'll be the end of the line for Earth. Two astronomers who suggested that our planet might escape the fireball as the dying sun expands have changed their minds: Earth will be engulfed and vaporized after all.

In 2002, Robert Smith of the University of Sussex in the U.K. and colleague Klaus-Peter Schroeder calculated that a strong solar wind would eventually reduce the sun's mass, weakening its gravitational hold on Earth. The planet would then drift outward far enough to remain intact.

But since then, Schroeder, now at the University of Guanajuato in Mexico, has revised the stellar mass-loss formula that astrophysicists have been using for decades. The new formula shows that the sun's gravity will be stronger than he and Smith thought. The two have also refined their calculations to take into account the way Earth's gravity causes the sun to bulge slightly toward it. This tidal effect slows Earth's course so that it sinks toward the sun. The results spell ultimate doom.

None of this matters much to human beings: Increased solar radiation will turn Earth into a lifeless rock about a billion years from now.



Out to get us.



In Print

A PACK OF ... ? A French surgeon has used his knowledge of “wolf children” to help expose the latest fabricated autobiography to rock the publishing world. Misha Defonseca, now living in Massachusetts, had claimed to be a Jewish girl from Belgium who lived with wolves during a part of her journey to Ukraine and back during World War II in a futile search for her deported parents. Published 11 years ago, *Misha: A Mémoire of the Holocaust Years* was turned into a feature film that premiered this year in France.

Her book drew the ire of Serge Aroles, who last year published a book debunking legends of children being raised by wolves. Although it's theoretically possible that a “pseudo-pregnant” female wolf would nourish a human newborn, Aroles says, a pack would never adopt a child of age 7. And he adds that certain details, such as a female wolf reprimanding Misha for urinating like a male, with one leg raised, were “just ridiculous.” Aroles also discovered that Defonseca, whose real name is Monique De Wael and who was born into a Catholic family, attended school during the years she claimed to have made the trip.

After Aroles published a number of online articles attacking the book and Belgian newspapers started investigating, Defonseca admitted to the hoax in a statement on 29 February. She asked forgiveness but said the story “has been my reality.”

IN MEMORIAM

CLOSURE. Airmen John Lubben, Albert Fogue, and Charles Spiegel disappeared in 1944 after their U.S. fighter plane went down during a bombing mission near Cologne, Germany. Thanks to efforts by forensic scientists, the men's remains will be buried with honors next month at Arlington National Cemetery near Washington, D.C.

Although the remains were discovered in Germany in 1975 and recognized as American servicemen, nothing suggested the men's identities. They were buried among nearly 800 unknown soldiers in a Belgium cemetery. Their story surfaced in 2003, when a group of German hobbyists unearthed American plane wreckage near the original location of the remains and notified the Joint POW/MIA Accounting Command (JPAC) in Hawaii. Forensic anthropologists at JPAC and scientists at the Department of Defense DNA Registry in Washington, D.C., identified the men by matching their fillings to their dental charts and comparing their DNA to that collected from family members.

JPAC issues about 100 identifications every year, some of them dating back to the American Civil War, and the DNA lab processes more than 800 samples a year to help identify casualties of past and current conflicts. James Canik, deputy director of the lab, says modern molecular methods help solve cases that would have been impossible a decade ago. “When we're able to go back to [the families] and provide them an answer after so many years, it's extremely rewarding,” he says.

DEATHS

BRIDGE. When China decided to look West in the late 1970s, hundreds of science

major clamored to get into Ph.D. programs in the United States—but they had no idea how to enroll, and U.S. universities had no idea how to assess them. Cornell University geneticist Ray Wu played matchmaker, creating the China-United States Biochemistry Examination and Application program, which vouched for students who did not have access to GRE or TOEFL exams.

Last month, Wu, 79, passed away in Ithaca, New York. To honor his memory, the Chinese Biological Investigators Society (CBIS) hopes to create a foundation that will continue his legacy of strengthening ties between Asia and the West.

Chinese biologists speak in awe of Wu, a soft-spoken émigré whose father, Wu Hsien,

is considered the founder of biochemistry in China. Wu came to the United States in the waning days of China's civil war in 1948 to

pursue doctoral studies. He made his own scientific mark in 1970, when he developed the primer-extension approach for determining nucleotide sequences. Frederick Sanger improved on the method and won a Nobel Prize. Later in his career, Wu worked on transgenic rice.

While mourning Wu, leaders of CBIS are drawing up plans for a Ray Wu Fund. One element of the fund will be prizes to outstanding graduate students of

any nationality working in Asia. And an October symposium in Ithaca, originally intended to celebrate Wu's 80th birthday, will now be a memorial event.



<< Two Cultures

THINK ABOUT IT. A statue of an upside-down figure, evocative of flames, is a new reminder of the value—and the price—of free thought. The memorial to Giordano Bruno, who was born 450 years ago, was unveiled on 2 March in Berlin's Potsdamer Platz station. The statue commemorates the one-time Dominican monk who was burned at the stake in 1600, primarily for denying the divinity of Christ, but also for supporting the Copernican model of the solar system and for his idea that the universe contained a multitude of suns and Earth-like worlds. The sculpture, by Alexander Polzin, is symbolic of the way scientific discoveries and revolutions happen, “by turning a worldview on its head,” says Jürgen Renn of the Max Planck Institute for the History of Science in Berlin, which hosted an associated symposium this week on Bruno. The 6-meter-high sculpture is designed to “irritate” passersby into reflecting on the role of human reason in making the world a better place, says Ernst Salcher of the Giordano Bruno Foundation, which helped fund the project.



Congress's new
physicist-politician

1470

Xenopus rules alarm
U.K. biologists

1472

SCIENTIFIC MISCONDUCT

Two Papers From Korean Lab Found To Lack 'Scientific Truth'

An investigation by a prominent South Korean university has revealed that two papers by its researchers "do not contain any scientific truth." Both will likely be retracted by the journals in which they appeared, *Science* and *Nature Chemical Biology*. The papers describe a new way to identify drug targets by tracking protein movements in living cells. Their well-known senior author, Tae Kook Kim, studied in the United States and founded a company in Daejeon, South Korea, in 2004 to commercialize the findings.

Kim, in a brief e-mail to a *Science* reporter, said he would cooperate with the investigation "to open the truth very soon." He added that a "certain party has twisted this current situation to take an advantage of it."

An inquiry by the Korea Advanced Institute of Science and Technology (KAIST) in Daejeon, where Kim is a faculty member, is not yet complete. But Gyun Min Lee, chair of KAIST's Department of Biological Sciences and head of the internal investigation committee, informed *Science* by e-mail that "our initial investigative results are strong enough to convince us that the two papers do not contain any scientific truth." Another member of the committee, Yeon-Soo Seo, a biochemist at KAIST, declines to say precisely what it has found until the investigation is complete.

KAIST launched the inquiry after scientists at Kim's company, CGK Co., had difficulty coaxing the technology, called magnetism-based interaction capture (MAGIC), to work. One co-author of both papers and a former Ph.D. student in Kim's lab, Yong-Weon Yi, contacted *Science* and *Nature Chemical Biology* in December to ask that his name be removed from the papers. The journals then quietly began asking questions of their own. *Science* was reassured by Kim that "he didn't think there were prob-

lems with the paper," says Katrina Kelner, *Science*'s deputy editor for biological sciences. The journals' inquiries were continuing when KAIST announced its preliminary findings on 29 February.



Too good to be true? Papers describing a technique known as MAGIC for tracking proteins in a living cell have been called into question.

At CGK, which raised \$2.5 million from three Korean venture capitalists in 2006, concerns about the technology had run deep for months. "CGK has tried to reproduce the technology but in vain," the company said in a written response to questions from a reporter for *Science*. In mid-February, CEO Jin Hwan Kim informed KAIST officials of the company's difficulties with MAGIC, prompting KAIST to launch a departmental investigation the next day.

The school followed a protocol established by the Ministry of Science and Technology after a scandal over fraudulent stem cell publications by Seoul National University professor Woo Suk Hwang (*Science*, 19 May 2006, p. 980). Seo says investigators approached Tae Kook Kim, who couldn't provide notebooks or original data for the experiments. The task force interviewed members of Kim's team and met twice with Jaejoon Won, the first author of both studies. After the second interview, Seo says, Won sent a written statement to the committee that, Seo says, admitted "serious scientific misconduct in both papers." The committee reported its findings to KAIST President Nam Pyo Suh on 28 February. The school then suspended Kim and notified the journals. *Science* posted an Editorial Expression of Concern about the paper online on 3 March, and editors at *Science* and *Nature Chemical Biology* say they're hoping to hear from all the authors, or learn the outcome of the investigation, before retracting the papers.

Both journals say the papers garnered praise. "Reviewers were very enthusiastic," says Kelner. Notes Terry Sheppard, chief editor of *Nature Chemical Biology*: "The referee comments, I thought, were quite consistent with a favorable response."

But in response to a reporter's questions, CGK listed eight problems with MAGIC, in which magnetized nanoparticles are prodded to interact with proteins in cells. Among them were the type of magnetic nanoparticle used and the resolution of the microscope that authors say generated the published images. It "cannot produce the results shown," according to CGK.

Science began routine screening of images for manipulation in 2006, after publishing Kim's paper in July 2005. The *Nature* journals instituted image screening a few months after publishing their paper from Kim and his colleagues in the summer of 2006.

Kim's mentors were stunned to learn of KAIST's initial findings. "I certainly would never have expected this," says Tom Maniatis, a molecular biologist at Harvard University who supervised Kim's postdoc work in the mid-1990s. Maniatis described Kim as "an extremely hard worker, very efficient, very focused," echoing comments



from Kim's Ph.D. adviser, Robert Roeder of Rockefeller University in New York City. The work Kim performed in Maniatis's lab, on interferon genes, has held up, as have his experiments under Roeder.

After his postdoc, Kim landed a coveted fellowship at Harvard's Institute of Chemistry and Cell Biology, as well as a post with Seoul National University's Institute for Molecular Biology and Genetics. It was at this time that he began focusing on ways to identify drug targets. MAGIC does this by coating a magnetized nanoparticle with a small molecule of interest; the particle is then gulped up by a cell. Proteins are tagged one after another with fluorescent labels, and scientists apply a magnetic field to the cell. This forces the nanoparti-

cle to move, and researchers can observe whether a particular protein moves with it, indicating that the small molecule has bound to the protein. The 2006 *Nature Chemical Biology* paper showed that MAGIC could pick up two proteins that reset a cell's aging clock.

Although the MAGIC technology was eye-catching, it apparently wasn't hugely popular. "I don't know anybody who was using" MAGIC, says Brent Stockwell, a chemical biologist at Columbia University who knew Kim somewhat when the two were at Harvard. Stockwell recalls reading the *Science* paper and wondering whether it "seemed too good to be true." Because the technique melded different disciplines, he adds, "it would be hard to find anyone who

would have expertise in all" of them.

Over the past year, Kim had become embroiled in a patent dispute between KAIST and CGK over rights to MAGIC. Last July, KAIST suspended Kim for 6 months for allegedly bypassing the school's normal patent procedures, and Kim had been aggressively job-hunting, seeking references from Roeder and Maniatis. "I must have sent out 30 letters in the last month," Maniatis says.

Kim was still listed as CGK's "scientific director and founder" on the company's Web site this week. But CGK says it has developed an alternative to the MAGIC technique and hopes to build a business around it.

—JENNIFER COUZIN AND DENNIS NORMILE

BIOMEDICAL RESEARCH

\$300 Million in Private Money for New Investigators

Concerned that young investigators are barely treading water, the largest U.S. biomedical research charity this week unveiled a plan to throw a lifeline to some of them. The Howard Hughes Medical Institute (HHMI) in Chevy Chase, Maryland, will spend \$300 million over 6 years to support researchers it considers promising but who are struggling to obtain their first independent federal research grant. Hughes aims to fund up to 70 this year.

HHMI President Thomas Cech sees the new program as an emergency effort to help young scientists who have been disproportionately affected by 5 years of flat budgets at the National Institutes of Health (NIH) in Bethesda, Maryland. As success rates for NIH research grants scrape a historical low, Cech and other leaders worry that many new investigators are abandoning academic research. Underscoring the community's concerns, a Senate committee this week held a hearing to look into the "Broken Pipeline" in life sciences—the title of a report* released the same day by a coalition of universities describing the difficulties faced by 12 young investigators.



Pipeline problem. HHMI President Thomas Cech worries about an academic brain drain; Ohio State's Jill Rafael-Fortney testified about challenges facing new investigators.

HHMI plans to help scientists with a lab and 2 to 6 years' experience as assistant professors—people who typically have an early career award but currently have only an 18% chance of being awarded their first basic NIH R01 research grant. Cech says that when he visits top research universities, "the early faculty are struggling and depressed." The one-time, 6-year appointments will average about \$700,000 per year—twice the size of an NIH R01 grant—split between research, salary, and institutional support.

Hughes's rescue is a drop in the bucket compared with the challenge NIH perceives. NIH aims to give R01 awards to the historic average of 1500 new investigators

annually, starting last year, up from 1353 in 2006. Other NIH programs have geared up: Pathway to Independence will give out about 170 awards this year for training and research, and New Innovators will award up to 24 5-year grants to new investigators for high-impact research.

Young investigators say help is badly needed. Human geneticist Jill Rafael-Fortney of Ohio State University, Columbus, who testified this week before the Senate Health, Education,

Labor, and Pensions Committee, is waiting for a renewal of her first R01 or a new grant to study a muscular dystrophy gene that also plays a role in heart failure. The second proposal has gotten "great scores, and it's just missing the pay line," or cutoff, for funding, she says. If her R01 applications fail, and a multi-PI grant she's on doesn't come through, it will mean a 50% cut in her salary. "I'd probably have to leave my job," she says.

Hughes expects to fund at least one more competition 3 years from now to help young people "through this rough period," Cech says. "We hope we won't have to do it any more than twice." —JOCELYN KAISER

* www.brokenpipeline.org

U.S. CONGRESS

Physicist Wins Open Illinois Seat

Developmental biologist Donald Brown of the Carnegie Institution of Washington in Baltimore, Maryland, doesn't get many fundraising calls from candidates for Congress. But when experimental physicist Bill Foster called Brown in January, he gave the Illinois Democrat his ear, and soon after, \$1000.



Bigger science. Bill Foster (center) will join two other physicists in Congress.

Last weekend, Foster, a former researcher at Fermi National Accelerator Laboratory (Fermilab) in Batavia, Illinois, won a special election to fill the seat held by former House Majority Leader Dennis Hastert in a race that attracted national attention. And hundreds of scientists around the country gave their time and money to help put Foster over the top.

"The combination of my scientific and business backgrounds ... spells problem solver," says Foster, who ranks the economic downturn and pulling U.S. troops out of Iraq as his highest priorities and says more funding for basic science is a problem he also wants to tackle. "Bill's background as a scientist not from Washington helped him," says his campaign manager, Tom Bowen.

During his 22-year career at Fermilab, Foster, 53, also was part-owner of a prosperous lighting firm. Unhappy with the cancellation of a particle physics project he led and bitten by the political bug, Foster left the lab in 2006 to work on the successful congressional campaign of another neophyte, Democrat Patrick Murphy, for a seat in suburban Philadelphia. Calling himself the campaign physicist, Foster learned the ropes—and devised software to organize the efforts of volunteers knocking on some 140,000 doors.

Foster's lab ties helped him jump-start his own campaign. Nobelist and former Fermilab director Leon Lederman introduced Foster to Chicago power brokers and assembled a list of 28 Nobelists who publicly endorsed Foster. After work hours, Fermilab physicist and county alderman James Volk recruited colleagues to get involved. Lab engineer Thomas Peterson told voters in a campaign commercial

that Foster "brings people together." Volk and others braved the bitter weather "to collect signatures and knock on doors to get out the vote," he says.

The community also dug into its pockets, giving roughly \$160,000 according to campaign documents. Amounts ranged from \$250 from former Department of Energy official and Massachusetts Institute of Technology physicist Mildred Dresselhaus to \$2000 from Fermilab Director Pier Oddone. Foster was his own biggest contributor, spending \$1.4 million, and the Democratic National Committee put in a similar amount to help Foster capture what had been a traditional Republican stronghold.

Brown says he's proud of his role in helping Foster win the Democratic primary last month and then defeat Republican Jim Oberweis on 8 March by a margin of 53% to 47%. "It's a great omen for the Democrats," he says. "This Administration has been terrible for science."

Despite being the newest member of Congress, Foster won't be able to leave the campaign trail. That's because he faces Oberweis again on 4 November in a race for a full 2-year term. And scientists such as Fermilab physicist Michael Church, who contributed \$400, say Foster can count on their support again. "It'll be a tough election, and he'd need all the help he can get," says Church. **—ELI KINTISCH**

SCIENTIFIC WORKPLACE

Indian Government Offers Helping Hand to Women Scientists

NEW DELHI—In 2000, when Vijayalakshmi Ravindranath was appointed director of the National Brain Research Centre in Gurgaon, the neuroscientist made history. Ravindranath became the first woman to lead any of the 65 institutes under India's Ministry of Science and Technology—and today she is one of only two women who have broken the ministry's glass ceiling. When it comes to promoting women scientists in India, she says, "our record is dismal."

It may be surprising that women scientists are struggling in the nation that elected Indira Gandhi prime minister in 1966. But at a conference in New Delhi to mark International Women's Day on 8 March, more than 1000 scientists spoke of barriers to advancement and debated how to attract more women into research careers. At the meeting, science minister Kapil Sibal announced what he calls

"fledgling steps to ... empower women to have their rightful role in science," including new regulations to allow women with young children to work more flexible hours.

The statistics are sobering. Although women have earned 37% of all science Ph.D.s awarded by Indian institutions, they hold fewer than 15% of science faculty positions. Out of India's 114,000 or so government scientists, fewer than 16,000 are women. Only one of the Indian Council of Agricultural Research's 175 institutes has a female director: the National Research Centre for Women in Agriculture in Bhubaneswar.

The hurdles women scientists face are not unique to India. "Women are unable to cope with the triple burden of home, work, and societal prejudices, including gender-related nepotism," says Mahtab S. Bamji, a nutritionist with the Dangoria Charitable

Trust in Hyderabad and chair of a science ministry task force assessing the status of women in science in India. The panel will release its findings later this year.

At the conference, Sibal announced that his ministry's 65 institutions would provide flexible working hours for women scientists with children younger than 3 and establish crèches in all institutions. In addition, Sibal said, some of the 33 female members of the Indian National Science Academy (which has a fellowship of 774) will receive research grants of up to \$25,000 a year over 5 years.

Women scientists applaud the measures. But in the long run, as India's first woman president, Pratibha D. Patil, argued at the conference, gender parity will come only when discrimination against women has stopped. **—PALLAVA BAGLA**

Dinnertime. Unprotected by DEET, this man's arm draws hungry mosquitoes.



NEUROSCIENCE

Hiding From Biting Insects in Plain Scent

For backpackers, the insect repellent DEET makes hikes bearable by protecting against mosquitoes and other blood-sucking pests. For people who live in malaria-prone areas, it can be a lifesaver. Although people have been spraying and dabbing DEET on for more than 50 years, nobody knew exactly how it works until now. Researchers report online in *Science* this week (www.sciencemag.org/cgi/content/abstract/1153121) that instead of driving away biting bugs, DEET actually conceals us from them. The compound dulls the insects' sensitivity to certain body odors, explains co-author Leslie Vosshall, a molecular neurobiologist at Rockefeller University in New York City.

Although DEET's insect-deterrent abilities are formidable, it can be toxic, especially to infants. Unraveling how DEET fends off attacks "is a first step" toward developing more specific alternatives, says neuroscientist Richard Axel of Columbia University, who wasn't involved in the research. The work "affords the possibility of rational approaches to confounding the olfactory system of insects." But first, more behavioral studies of mosquitoes are needed, cautions medical entomologist Willem Takken of Wageningen University in the Netherlands, as Vosshall's behavioral experiments focused on fruit flies.

To female mosquitoes, our body odors mean food. These insects sniff out lactic acid in our sweat and carbon dioxide and 1-octen-3-ol alcohol in our breath. "It's wildly attractive to mosquitoes," says Vosshall.

She and her colleagues netted key evidence about DEET's mode of action from fruit flies, whose neurobiology is better understood than that of mosquitoes. For these behavioral studies, the researchers gave the flies a chance to enter either of two small

vials. If both tubes held food, flies piled into them in about equal numbers. If one tube also contained DEET at the 10% concentration found in many sprays, the insects preferred the DEET-free vial. However, when the choice was between an empty tube and an empty tube dosed with 10% DEET, the flies made no distinction, suggesting that they weren't just shunning the compound. With DEET wafting around, they apparently couldn't smell the food, says Vosshall.

To bolster this suspicion, her team took a close look at fly odor receptors. The team engineered frog eggs to make these receptors, a technique that effectively "can turn a frog egg into a fly nose," says Vosshall. Exposure to certain odor molecules activated the receptors and triggered a surge of positive ions into the egg. But DEET reduced the flow, indicating that the receptors are less responsive in DEET's presence. The compound hides food scents by blunting the sensitivity of specific odor receptors, the researchers conclude.

But that conclusion is premature, says Takken, until the researchers determine whether mosquitoes react the same way as do fruit flies in the tube-choice tests.

However, Vosshall and colleagues did show that their results held true for mosquito receptors and neurons. Using the frog-egg method, the team discovered that DEET suppresses three body-odor-detecting receptors from *Anopheles* mosquitoes, which transmit malaria. They also saw a dampening of electrical activity in mosquito neurons that typically respond to 1-octen-3-ol, confirming that DEET disrupts the insects' perception of some attractive smells.

Calling DEET a "repellent" is a misnomer, says Vosshall: "It doesn't smell bad to insects; it masks or inhibits their ability to smell you."

—MITCH LESLIE

More Wins for Wisconsin

The U.S. Patent and Trademark Office (PTO) has upheld two key patents, on primate and human embryonic stem cells, held by the Wisconsin Alumni Research Foundation, giving WARF a solid victory in its hotly contested patent battle. Last month, PTO upheld a third patent, on stem cell culture techniques (*Science*, 7 March, p. 1323). The WARF patents, granted in 1995, 1998, and 2001, were challenged 2 years ago by two nonprofits that said the work was obvious and therefore unpatentable. The patent office disagreed, although it narrowed the patents slightly. "We applaud this final decision on our two most important base stem cell patents," said the foundation's Carl Gulbrandsen.

—ELI KINTISCH

New Panel Tackles Boston Peril

Responding to state and public concerns about the safety of a new biodefense lab being built by Boston University, the National Institutes of Health (NIH) has set up a blue-ribbon panel to provide an independent assessment of the \$128 million facility. "We will thoroughly examine any potential risks to the community associated with this project," says Anthony Fauci, who directs NIH's National Institute of Allergy and Infectious Diseases, which is funding the lab. The panel, which held its first meeting this week in Boston, is chaired by Adel Mahmoud, a Princeton University biologist and former president of Merck Vaccines. The lab, which will handle highly toxic and infectious biological materials, has drawn widespread criticism for its location in a densely populated and mostly minority area of downtown Boston. —ANDREW LAWLER

Research Funds Granted

BEIJING—China's premier research labs have many advances to their credit, including super-hybrid rice varieties, new dinosaur species, and instruments on the Chang'e lunar spacecraft. Now the government is showing its appreciation with a huge cash injection. Last week, the science and finance ministries announced the creation of a special fund for China's 220 key state laboratories. The fund will hand the labs \$280 million in 2008, primarily for research materials and new equipment. The money is expected to liberate key-lab scientists from the drudgery of grantsmanship. "Our scientists have wasted a lot of time applying for funds," says Chen Changjie of Beijing's Institute of Biophysics, which has key labs on biological macromolecules and brain and cognitive science. Now, Chen says, China's top scientists "will have more time to do research." —LI JIAO

DEVELOPMENTAL BIOLOGY

Proposed Frog Ban Makes a Splash

A proposal to ban the sale of African clawed toads in the United Kingdom has that country's developmental biologists worried about the availability of their favorite research animal—and conducting studies to surmount the reason for the ban. In November, the U.K.



Environmental menace? African clawed toads can spread disease.

Department for Environment, Food and Rural Affairs (Defra) issued a list of potentially harmful non-native species that it wants to ban from sale or trade. To the surprise of scientists who work with the animal, the toad, *Xenopus laevis*, was on the list. Defra says the animals can carry the chytrid fungus blamed

for mass amphibian die-offs around the world. It also noted that escaped African clawed toads can compete with—and sometimes eat—native amphibians.

The U.K. biomedical charity The Wellcome Trust and other scientific groups quickly lodged protests and filed briefs against the proposed ban. Defra's list "came out of left field," says developmental biologist Matthew Guille of the European *Xenopus* Resource Centre in Portsmouth. The two sides are likely to settle on a new licensing scheme that would limit, rather than ban, sale of the animals. But the issue has made biologists more aware of potential dangers their toad colonies pose to the environment.

For decades, developmental biologists have prized *Xenopus* for its large, robust embryos that make it possible to observe and manipulate key stages of development. In recent years, the frogs have become even more valuable to researchers, as scientists have learned to create mutants and interrupt the function of specific frog

genes. A ban on the sale of the animals would mean that researchers would have to keep their own breeding frogs on hand, making it much more difficult and expensive to manage colonies, Guille says.

Defra admits being surprised by the scientists' reaction. "We didn't realize it was quite such a big deal," says Matthew Ashton, who is coordinating the agency's response.

The scientists support Defra's aims, if not its ban. "We are biologists. I have no wish—nor do my colleagues—to be the people who are responsible for the die-off of a whole lot of native amphibians," Guille says. The *Xenopus* research community already had relatively strict standards in place to prevent wild release of their animals, he says, but in response to the proposed ban, researchers have started to test for the fungal disease, which doesn't make *X. laevis* obviously ill. Guille says that so far, levels in laboratory frog colonies seem to be low. Although there is no evidence that the fungus compromises studies by affecting *Xenopus* development, "we'd all feel a lot better if we knew we had an infection-free colony," he says. There is no recognized treatment, he says, but the fungus seems to disappear on its own when animals are kept at low density.

Ashton says Defra's final decision on the *Xenopus* ban is expected in October.

—GRETCHEN VOGEL

EDUCATION

Spanish Scientists Question Accreditation Plan

Many in Spain's scientific community are lobbying the country to change a new accreditation system for professors, complaining that it emphasizes nonresearch activities such as patenting, consulting, university management, and work in governmental and state agency offices. In a letter last month to the Spanish Minister of Education and Science, endorsed by more than 760 Spanish scientists, Joaquín Marro Borau, a statistical physicist at the University of Granada, expressed the concern that "many of the best scientists ... would never be able to become full professors in Spain with this marking scale" and that many competent scientists will "autoexclude themselves from the process."

Some Spanish officials have questioned the political motives of the letter—at least one signer had connections to the opposition party that lost last week's elections—and

others argue that scientists are rushing to judgment. "When accreditations start to be given out to scientists and they see that someone of their profile gets it, they are going to calm down," says Félix Pérez Martínez, coordinator of faculty evaluation for Spain's National Agency for Quality Assessment and Accreditation (ANECA).

Marro denies any political motive to the letter. And the debate over accreditation dates back to 2001 when Spain, in an attempt to bypass the old-boy network, mandated that aspiring professors first get their competences accredited. But everyone soon agreed that that process was flawed.

The new accreditation scheme replaces oral examinations with faster, electronic applications. But a complex scoring system, released in January, has come under fire for giving too much credit to nonresearch-related activities—an applicant to full

professorship must get 80 out of 100 points, although research can provide a maximum of 55 points. But the authors "have misinterpreted the importance of this marking scale," says Spanish Secretary of State for Universities and Research Miguel Ángel Quintanilla Fisac. Evaluators, adds Pérez, can give scientists points for management activities such as coordinating European research projects and sitting on editorial boards for journals.

All agree that the revised application procedure remains cumbersome. Applicants must back each achievement with written proof, for example. "They've made it so bureaucratic, difficult to comply with, and confused [that] it seems unbearable," says Maria Àngels Garcia Bach, an associate professor at the University of Barcelona.

ANECA says that it's open to adjusting the process further. "I believe that we still need to improve and ask the ANECA to make [the system] more agile," Quintanilla says.

—ELISABETH PAIN

Elisabeth Pain is a contributing editor for *Science* Careers.

INNOVATION

U.K.'s Royal Society Ventures Into Funding Start-Up Companies

The Royal Society, the United Kingdom's academy of science, is taking the unusual step of getting into the venture capital business. Last month, it announced the creation of an Enterprise Fund, with the aim of funneling money into start-up companies seeking to commercialize the fruits of academic research. The society says that although it does much to support science and teaching, as well as contributing to policy debates, it could be more active in applying science for public benefit—either by simply boosting the U.K. economy or tackling problems such as carbon capture, renewable energy, and water purification. “If there are difficulties getting science into the marketplace, the society has a role helping with that,” says Andrew Mackintosh, a chemical physicist-turned-businessman who has been brought in to run the fund.

Although the society's plan has largely earned plaudits, some worry that the institution could ultimately tarnish its famous name, and others argue that any money raised under the society's name should go to traditional research. Nick Dusic, head of the Campaign for Science and Engineering, a U.K. pressure group, welcomes the injection of start-up funding but says the society “mustn't lose its independence in advising government and disbursing funds.”

Some see the Enterprise Fund as a return to the society's roots, as many of its early activities were spurred by the business concerns of the day. “The society was a more commercially oriented organization back then,” says Peter Littlewood, head of physics at Cambridge University and a Fellow of the Royal Society. In the run-up to its 350th anniversary in 2010, the society is seeking to raise £100 million (\$200 million) for a variety of special projects, £20 million of which is destined for the Enterprise Fund. The society has so far raised £3 million for the fund from philanthropists, and Mackintosh says he hopes to begin making investments this year.

The equity investments may range from £250,000 to £2 million but would typically be about £500,000 and extend over longer time scales than are usual for venture capital funding. The fund will be commercially run and legally independent from the society, and any profits will be reinvested in the fund. The hope is that the society's network of grantees, as well as 1400 elected fellows, will be a profitable source of ideas for start-up com-

panies. An advisory board of fellows and some well-known venture capitalists will oversee the fund.

Medical start-ups in the United Kingdom are frequently given help by organizations such as the Wellcome Trust or cancer research charities, but “nonmedical research is more difficult. There's a severe shortage of early stage venture capital,” says Roger Brooke, a venture capitalist with the IP Group and one of the fund's early benefactors.

Mathematician Michael Atiyah, a former president of the Royal Society, says “this is a good thing for the Royal Society to be doing, well within the scope of an academic body.” The Royal Society of Edinburgh, a separate Scottish body of which Atiyah is now presi-



Bold venture? The venerable Royal Society wants to help scientists with good ideas turn them into commercial reality.

dent, has had a similar scheme in place for 10 years using government money, he says, and it's been “very successful.”

Not everyone is entirely convinced, however. Littlewood is skeptical that there is a problem finding financing for start-ups with a good idea. “This is yet another surge in a direction where we don't need much help,” he says. “I don't feel it's inappropriate, but is the Royal Society the best vehicle to do it? It's increasingly hard to get money just to do science.”

—DANIEL CLERY

Message From Mars Mottled

Planetary scientists who study Mars like NASA's plan to bring back samples (*Science*, 29 February, p. 1174). But they think the agency's plans to pay for it are unrealistic and ultimately bad for those who study the planet. A 29 February letter from the Mars Exploration Program Analysis Group to the NASA Advisory Council's Planetary Science Subcommittee warns that the plan would have “a devastating effect on the Mars program” by creating a 13-year gap in landings and causing “severe damage” to the next generation of scientists. The two-page letter urges NASA to conduct a sophisticated sample return rather than a limited mission that would pick up material gathered by the upcoming Mars Science Laboratory.

NASA's planned sample return “cannot happen” unless the agency doubles its current Mars spending levels, states the letter, signed by planetary scientist John Mustard of Brown University, who chairs the group. And a 2-year delay won't work, the letter adds, because the orbital trajectory between the two planets is unfavorable. Sean Solomon, a Carnegie Institution of Washington scientist who chairs the advisory subcommittee, says his panel also supports the more expensive and sophisticated sample-return mission.

—ANDREW LAWLER

Australian Reactor Idle

CANBERRA, AUSTRALIA—With the country's only nuclear reactor shut down, Australian nuclear scientists are hoping the authorities get them back in business soon. The Australian Nuclear Science and Technology Organisation (ANSTO) shut down the Open Pool Australian Light-Water (OPAL) reactor in Lucas Heights outside Sydney 7 months ago because of faulty uranium-laden fuel elements. ANSTO has been working to solve the problem, but the government regulator wants more details before it will clear the reactor to start up again, ending the shutdown.

The delay is hampering scientists who use OPAL to conduct neutron-scattering research to probe high-temperature superconductors and other advanced materials. Robert Robinson, head of the Bragg Institute in Lucas Heights, laments “the potential loss of users to overseas facilities and the impost on the users, particularly graduate students, who have either not been able to do their research or have had to fund travel to overseas neutron sources.”

—CHERYL JONES



Dueling Visions For a Hungry World

Sparks began to fly when scientists and activists against genetically modified crops came together to assess agricultural knowledge and the role of biotech in development

When economist Carl Pray heard about plans for the first international assessment of agricultural research, a gold standard sprang to mind: the Intergovernmental Panel on Climate Change (IPCC). But things didn't turn out the way he expected.

IPCC has been pivotal in proving that climate change is real and linking it to human activities. As an agricultural economist at Rutgers University who has worked in many poor countries, Pray is convinced that agricultural research—and genetic modification in particular—is key to fighting pervasive hunger, which will only worsen as the world's population soars to 9 billion during the next 50 years. He hoped that the new project, like IPCC, might be able to muster top experts and galvanize support for a field that had languished through the 1990s. Most encouraging was the leader: Robert Watson, who had earned high marks as a chair of IPCC.

By the time Pray had signed up and the International Assessment of Agricultural Science and Technology for Development (IAASTD) got under way in 2005 with funding from the United Nations, the World Bank, and several countries, the purview had broadened far beyond food production to include social justice and the environment. The overarching question, posed on the home page of IAASTD's Web site,* is a mouthful: "How can we reduce hunger and poverty, improve rural livelihoods, and facilitate equitable, environ-

mentally, socially and economically sustainable development through the generation, access to, and use of agricultural knowledge, science and technology?" Critics say this broad mandate made conflict inevitable and stunted the assessment's analytical rigor.

On several key issues, consensus proved elusive. Industry scientists and some academics—mainly agricultural economists and plant biologists—believe the assessment was "hijacked" by participants who oppose genetically modified (GM) crops and other common tools of industrial agriculture. Tensions peaked in October when Monsanto and Syngenta walked out of the assessment.

Many other participants, who include activists and social scientists, are pleased with

Key Findings of IAASTD

- Focus on "agroecological" strategies to address environmental issues.
- Create opportunities for poor farmers and rural laborers.
- Do more to involve women to advance toward sustainability and development goals.
- Integrate formal, traditional, and community-based knowledge.
- Create space for diverse voices and include social scientists in policy.

the outcome. They note that the voice and experience of small-scale farmers, particularly women, have finally been brought to the fore by the assessment. "It really deals with issues of power, influence, and benefits," says Marcia Ishii-Eiteman of the Pesticide Action Network North America in San Francisco, California. Toby Kiers, who studies sustainable agriculture at Vrije University in Amsterdam, the Netherlands, agrees. "For technology to be most effective, farmers must be at the center, influencing how it is developed, delivered, and managed," she says.

But as the drafts are being readied for a final meeting of governments next month, the camps for and against GM crops remain polarized with no resolution in sight. Industry groups, once considered key "stakeholders," are boycotting the meeting, and some participants suspect the dissent will undercut the assessment's impact. Says Piet van der Meer of the Public Research and Regulation Initiative in Delft, Netherlands: "There is a sense of having lost a wonderful opportunity."

Big tent

The origin of the \$12 million IAASTD dates to 2002, when a group of ag biotech companies asked the World Bank whether it recommended GM crops for developing countries. Watson, then the World Bank's chief scientist, suggested that the bank review the entire range of agricultural technologies and policies. Convinced that agricultural research should be considered in the context of the myriad factors

* www.agassessment.org

Green machines. Experts debated the role of tools, such as fertilizers, that have boosted yields.

that affect farmers, Watson and his team solicited suggestions about what to include in the assessment from some 800 stakeholders representing scientific and agriculture-related organizations around the world. "Right from day one, I wanted to make sure there was a wide range of views," Watson says. He formed an advisory bureau of 55 people, comprised of representatives from governments, consumer groups, industry, and nongovernmental organizations such as Greenpeace.

In theory, such an inclusive approach can be a plus, says William Clark, a science policy expert at Harvard University, conferring political legitimacy on an assessment. In practice, it can lead to logjams. Part of the tension over IAASTD reflects two competing worldviews of agriculture. Hans Herren, one of IAASTD's co-chairs, stresses the importance of recognizing the "multifunctionality" of agriculture. That is, in addition to producing food, farmers have other important roles, such as maintaining the landscape and cultural heritage. In contrast to this view, Herren says others "see agriculture as an industrial process, like making cars." That analogy doesn't offend Jonathan Gressel, a plant scientist with the Weizmann Institute of Science in Rehovot, Israel. "Producing food is the primary function of agriculture," says Gressel. "It should do it safely and [should be] compatible with the environment, just like GM [General Motors] should produce a safe and nonpolluting car—but they have to produce cars."

Many of the players weren't used to sitting at the same table with opponents. "We constantly had doubts about participating," says Benedikt Haerlin. At the time, he represented Greenpeace, which wants a ban on GM organisms (GMOs) in the environment and a phase-out of pesticides and a reduction of synthetic fertilizers. Greenpeace is not used to compromising, Haerlin concedes.

Molecular biologists and groups that employ them were also skeptical. Rodney Brown of Brigham Young University in Provo, Utah, who was then deputy undersecretary for research, education, and economics at the U.S. Department of Agriculture, worried that the scientific analyses might be swamped by nonscientific views. "Like it or not, not all input is equally valuable," Brown says. But almost all participants and observers interviewed by *Science* say that they had faith in Watson. "He has one of the finest, sharpest minds for assessments," says Walter Reid, now at the David and Lucile Packard Foundation in Los Altos, California, who led the Millennium Ecosystem

Assessment that examined the state of the world's ecosystems as of 2000 (*Science*, 1 April 2005, p. 41).

Watson—who is now at the University of East Anglia in Norwich, U.K., and chief science adviser to the U.K.'s Department for Environment, Food and Rural Affairs—was involved in several other projects, so much of the supervision of the assessment fell to his two co-chairs. Herren, a longtime advocate for sustainable agriculture, is the president of the Millennium Institute, a think tank in Arlington, Virginia. Adding perspective from a developing nation is energy specialist Judi Wakhungu of the African Centre for Technology Studies in Nairobi, Kenya. The pair convened meetings at which the advisory bureau picked the 400-odd authors and assigned them to groups to tackle the many topics, which included bioenergy and the role of women in agriculture, as well as five geographically focused subassessments.

A question of balance

It's perhaps not surprising that this massive effort encountered some snags. Some authors of the eight main chapters say that from the start, the process was disorganized, lacked clear direction, and suffered from turnover of participants and continual rehashing of drafts. It didn't help that the makeup of the teams was fairly rigid. Each chapter had to have the same number of men and women as authors. One of the two lead authors on each chapter had to be a woman, and one had to be from a developing country. "We wanted to make sure that everyone was represented equally,"

says Herren. Pray has a different take: "It was excruciatingly politically correct in some ways."

Some teams bogged down in conflicts about hot-button issues such as GM crops or trade liberalization, with various members charging each other with bias. One chapter, on how to help developing countries generate and adopt agricultural research, was canned entirely. "It was clear that the chapter was not going to fly," says Beverly McIntyre of the World Bank, who was a senior administrator on the project. She says the team did not have the right expertise for the task. And midway through, just before initial results were to be presented, the bureau decided to eliminate a major modeling exer-

cise. The International Food Policy Research Institute had raised about \$460,000 for the modeling, which would have provided insights to help policymakers compare the outcomes of four broad policy scenarios, such as futures with more free trade or green technologies. But Greenpeace's Haerlin and others objected that the models were not "transparent."

Conflict erupted in the review process as well, with some scientists and GM advocates complaining that their comments were not incorporated. To a certain degree, they may have been outmaneuvered; environmental groups set up a well-organized Web site[†] to funnel comments to chapter lead authors. They may have been outnumbered as well. Bureau member Emile Frison, director of Biodiversity International, a research organization working to conserve agricultural biodiversity, says it was difficult from the outset to engage the best scientists "There are probably multiple reasons, including the fact that they didn't see it as important," he says. One participant who asked not to be named puts some of the blame on Watson for not spending enough effort personally recruiting top scientists, as Reid did with



Scrutinized. Perhaps the hottest issue was the role of genetically modified crops in helping poor farmers and making agriculture more sustainable.

the Millennium Ecosystem Assessment.

The text of the ag assessment concedes the difficulty of resolving the differences to everyone's satisfaction: "One of the key findings of the IAASTD is that there are diverse and conflicting interpretations of past and current events, which need to be acknowledged and respected"—not exactly a helpful insight for policymakers trying to decide whether to approve the planting of GM crops, for example. "It devolved into 'I'm okay; you're okay,'" says Andrew McDonald, a crop and soil scientist at Cornell University.

[†] www.agassessment-watch.org/review.html

Critics such as Adrian Dubock, who participated while serving as Sygenta's head of biotech ventures, also complain that the summary and synthesis documents—which are all most people will read of the 2500-page draft assessment—are biased against GM crops and don't incorporate positive appraisals from some of the underlying chapters. Chapter 5, for example, concludes that both transgenic and conventional breeding will be needed to boost crop productivity during the next 50 years. Eliminate transgenic crops, the chapter notes, and "humanity would likely be more vulnerable to climatic and other shocks and to increased natural resource scarcity."

But the Global Summary for Decision Makers puts less emphasis on the potential benefits of GM crops, says Dubock, who resigned in protest, and others. In the synthesis report as well, the section on biotechnology tends to highlight the controversies about GMOs, such as "lingering doubts about the adequacy of efficacy and safety testing." The majority of the four authors of this synthesis chapter have experience in risk assessment and sociology but not in the agricultural industry or plant modification. Like all the other authors, they were picked by the bureau, which had relatively few representatives from industry. "The whole thing was incredibly stacked" against GM crops, claims Gressel. An IAASTD official disputes that charge and notes that Deborah Keith of Sygenta was slated to write the first draft but resigned. "I felt it was a hopeless cause," Keith says.

Pullout

Last October, the biotech companies voted with their feet and pulled out of the process. Their representatives won't attend the final meeting. "We can't endorse something that is fundamentally giving the wrong message," explains Keith Jones, manager of stewardship and sustainable agriculture at CropLife International, an industry trade group. He says that the report tends to overstate the potential of organic and "ecological" agriculture, which he doesn't think is a viable solution for boosting global agricultural productivity.

Watson, who says he was "extremely disappointed" in the companies' decision to withdraw, argues that instead of giving up, industry scientists ought to have played a bigger role. "They should have screamed and pushed," he says. But he also accepts some responsibility



Backbreaking. The assessment concludes that agricultural research needs to do more for poor farmers such as these indigenous wheat growers in Bolivia, by better managing pests and soils, for example.

for not keeping close enough tabs on their concerns. "It means I didn't succeed as director at keeping all the players at the table." Herren decries the 11th-hour move as well, pointing out that industry reps sat on the bureau, which vetted all of the authors. "You cannot come at the end, after all the meetings setting up the rules of the game, and say, 'I don't like it.'" By walking out, industry ended the dialogue, Herren says: "You can agree to disagree, but that can't happen when you slam the door."

Not all who were displeased left. In December, for instance, the Alliance Executive of the Consultative Group on International Agricultural Research (CGIAR)—the governing body of 15 publicly funded scientific centers around the world—wrote a private letter to the bureau about its concerns. "The reader would get a rather negative view about agricultural research in general," says Frison, the chair of the alliance. In the letter, which *Science* has obtained, CGIAR notes that the assessment chapters might undermine support for research. But CGIAR has remained in the assessment. "It is more constructive to make our points as a participant," says Frison, who also appreciates the assessment's emphasis on the importance of involving farmers in research.

At a final meeting in April, representatives from about 90 governments will gather in Johannesburg, South Africa, to decide whether to endorse the report. (They can accept, approve, or just note it.) Watson doesn't think that industry's pullout will lessen the impact of the report, which he hopes governments and international donors, for instance, will use to

guide their investments in agricultural research.

Robert Paarlberg of Wellesley College in Massachusetts, author of *Starved for Science: How Biotechnology Is Being Kept Out of Africa*, is skeptical. "It's a document that has much less scientific credibility" than does IPCC, he says. By being so inclusive, it ended up more a collection of opinions than an incisive summary of the scientific literature. And because its scope is so broad, the assessment doesn't offer targeted analyses for particular problems. "You end up with [platitudes] such as 'Small farmers need to be supported,'" says Emmy Simmons, an agricultural development consultant who retired from the U.S. Agency for International Development in 2005.

Watson remains sanguine. He and others think the assessment will bring more attention to the plight of the rural poor and the chronic underinvestment in agricultural research. "If we can stimulate a debate," he says—for instance, about the degree to which agricultural science is meeting the needs of the poor and whether everyone gains from free trade—"then it's a success." Nor does he have any regrets about throwing the doors wide open. "I always knew it was a social experiment," he says.

As for Pray, he has mixed feelings. "Halfway through this painful exercise, I thought, 'If [Watson] can pull this off, it will be great,'" he says, "but we couldn't come to consensus. Now Greenpeace and Monsanto continue to beat each other up." Meanwhile, he says, neither the environment nor the poor are getting the agricultural research they deserve.

—ERIK STOKSTAD

GRADUATE TRAINING

NSF Fellowships Called Powerful Tool for Building the Pipeline

A prestigious program is slated for a big increase next year thanks in part to recent data showing its ability to draw talented students into science

A fellowship program that has nurtured the early careers of Nobel prizewinners and a co-founder of Google has quietly been gaining political favor in Washington. Last month, the Bush Administration proposed a sharp increase in the number of Graduate Research Fellowships (GRFs) awarded by the U.S. National Science Foundation (NSF). And Democratic presidential hopeful Senator Hillary Clinton (D-NY) has called for its expansion as part of her "innovation agenda."

The GRF's rise to prominence has been propelled in part by the work of Harvard University labor economist Richard Freeman, who calls the fellowship "the U.S.'s premier award for science and engineering graduate students." In a 2005 study, Freeman found that the GRF—which provides a 3-year stipend to U.S. citizens or permanent residents for graduate study at the university of their choice—is a powerful mechanism for luring talented U.S. students into science (www.nber.org; working paper 11623). He argues that NSF should triple the number of GRF awards (from roughly 1000 to 3000 a year) and increase the amount of the scholarship (now \$30,000) by \$10,000. Although NSF has responded with a more modest proposal—a 32% budget increase that could mean as many as 700 new awards in 2009—agency officials are happy to disseminate Freeman's message to make their case.

Created in 1952, the merit-based program funds all science and engineering disciplines proportionately—that is, it selects the same percentage of applicants in each field that NSF supports. In the early years, the physical sciences and engineering accounted for the bulk of the awards. Those fields now make up about half the recipients, with the life sciences comprising another quarter and the social and behavioral scientists about one-sixth.

The program helped nurture a post-Sputnik generation of researchers. Its 43,000 alumni include more than 20 Nobelists; a more recent recipient is Sergey Brin, a co-founder of

Google. But by the 1990s, it had fallen on hard times. Stipend levels had failed to keep up with those of other fellowships, and the number of awards had remained constant for decades despite a steady rise in the number of undergraduates majoring in science and engineering. The shrinking ratio, from 5.4 awards per 1000 science and engineering majors in its first few years to 2.2 earlier this decade (see graph), "sent a very bad signal to students that science wasn't important," says Freeman.

Students responded by avoiding the pro-

Analyzing 50 years of data, Freeman calculated that increasing the number of awards substantially would have only a slight impact on quality while resulting in an influx of students into science and engineering. A survey of 1800 Harvard undergraduates similarly suggested that talented students would be lured into science if NSF gave out more fellowships to support their graduate studies. In particular, some 40% of students said that they would opt for graduate school in science if given "a national fellowship"; only 18% already had such plans.

Freeman's analysis soon gained adherents. The Brookings Institution's Hamilton Project to foster economic growth has promoted it (www.hamiltonproject.org; policy brief 2006-09), and Clinton has embraced the tripling of awards and a 33% boost in the stipend, asserting that the fellowships are "the key financial resource for [U.S.] science and engineering graduate students."

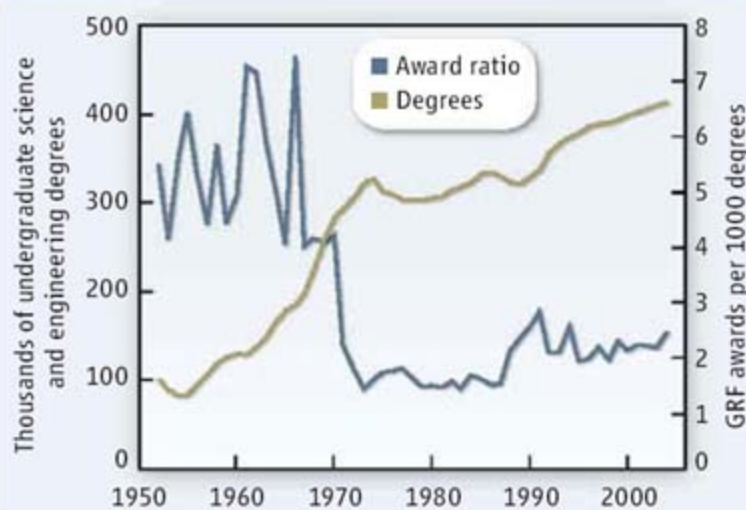
It also struck a chord at NSF. "He put the pieces together," says Carol Stoel, acting head of graduate education at NSF. The agency expects to make about 825 awards this year, says program director William Hahn, down from 925 in 2007. (For the past 2 decades, the annual number has fluctuated between 800 and 1000.) Its 2009 budget request of \$117 million for the education directorate, up from the current \$88 million, would raise the current 10% success rate for applicants. Because the fellowships run for 3 years, however, Congress would need to sustain the increase for several years to avoid sending the program on a roller-coaster ride.

As the biggest single increase in NSF's education budget, the request for the GRF program will attract close scrutiny from Congress. It must compete against programs serving different populations that enjoy strong legislative support, in particular those for minorities, undergraduates, and precollege science education.

"It deserves serious consideration because NSF feels so strongly about it," Representative Alan Mollohan (D-WV), chair of the House panel that controls NSF's budget, explained after a 27 February hearing in which NSF made its case. "Is it the fastest way to fill the pipeline? I don't know. And I don't want to prejudge [the request]. But there are a lot of things we need to do to improve science and math education."

—JEFFREY MERVIS

Not Keeping Pace



Negative numbers. A declining ratio of awards to degrees sends students "a bad signal," says Harvard's Richard Freeman.

gram. Hoping to turn things around, NSF began raising the annual stipend, which stood at \$15,000 in 1999, until it reached \$30,000 in 2004. (Universities also receive \$10,500 as an institutional allowance.) The number of applicants nearly doubled, Freeman found, to almost 9000 in 2004, and the overall quality of the applicant pool also rose. In other words, he says, the fellowship's prestige caused more of the "best and brightest" seniors to consider graduate degrees in science and engineering.

It certainly opens doors at Harvard, Freeman explains. "My department has a policy that if you get a GRF, you're admitted," he says. "Absolutely. Because you're not costing us anything. And we know NSF runs a pretty tight ship."



MEDICINE

Blood-Matching Goes Genetic

Hoping to prevent adverse transfusion reactions and save lives, European researchers are lobbying to replace serology-based blood typing with matching based on DNA tests

Two decades ago, a 20-year-old man entered Duke University hospital for a hip replacement. He had sickle cell disease and because he was anemic, he needed a blood transfusion before the operation. As with most sickle cell patients, he had received transfusions before. But on this occasion, his body rejected the blood.

The doctors tried again and again with different blood, but the man's immune system rebelled against every new transfusion, generating antibodies that killed the new blood cells and some of his own in the process. "He reacted to everything," says Wendell Rosse, a hematologist at Duke at the time. The man's anemia continued to worsen, and 2 weeks later he died. "There wasn't anything we could do," Rosse says.

Following previous transfusions, the man had become alloimmunized, meaning that his body made antibodies to donated blood. He had so many transfused cells mingling with his own and had made so many antibodies that physicians could no longer identify his original blood type nor find suitable blood. Although this rare case of fatal alloimmunization happened more than 20 years ago, Rosse says he's not sure whether the man would have survived today, as there have been few advances in blood-matching tests in hospitals.

Some hematologists believe genetic testing offers a solution. For more than 100 years,

blood matching, also known as typing, has relied on serology, the identification of surface proteins and carbohydrates, called antigens, on red blood cells. Genotyping looks at the genes that determine these antigens, and some scientists say genetic techniques can more closely match blood between donors and recipients, preventing alloimmunization and other immune-related blood reactions.

Now, these claims are being put to the test. France, for one, will begin genotyping a portion of blood donors by the end of April to determine if a switch from serology is scientifically justified and practical. In addition, Canada has begun using genotyp-



Old-fashioned. Serology tests like this one have been used to match blood samples for more than a century.

ing as a screening tool to identify donors with rare blood types that are hard to find in emergency situations.

"The argument that we will have 100% perfect blood for every patient will never be the case," says Neil Avent of the University of the West of England (UWE) in Bristol, who has led a European research consortium exploring blood genotyping. "But when you have all the information available, you can find the best blood."

Matchmaking

Finding the best blood seems simple. Two antigens on red blood cells determine the major blood group—A, B, AB, or O. A person with type A naturally harbors antibodies to B antigens and vice versa. Give type-B blood to a patient with type A, and antibody reactions will cause blood cells to clump together, sometimes with fatal results. AB people don't have antibodies to either antigen, making it easier to find blood for them. But O people have antibodies to both A and B, as their blood cells have neither antigen, so they can only receive O blood.

But there's more to blood than A and B antigens. Physicians also routinely test for D, an antigen in the Rhesus (Rh) blood group that can evoke a strong immune response. They try never to give RhD+ blood to an RhD- person.

The possibilities get even more complicated. Avent, who also directs UWE's Centre for Research in Biomedicine, says there are 29 known blood groups, which are determined by about 200 antigens. A person's full blood type might read: AB, D+, M+, N+, K-, Lea+, and so on.

A simple antibody test can identify a donor's or patient's ABO blood type, and in theory, similar tests could detect the other so-called minor blood groups. But the reagents to identify some of these antigens are costly, not reliable, or simply not available. For patients who receive only one or two transfusions during their lifetimes, mismatches in the minor blood groups pose no obvious problems. Yet people who receive multiple transfusions—including those with sickle cell disease, hemophilia, or leukemia—can develop antibodies to the minor blood group antigens, which is what happened in the Duke University patient. In some cases, the transfused blood can cause acute or delayed hemolytic reactions. The Public Health Agency of Canada, which keeps statistics on such reactions, estimates that as many as 1 in 12,000 transfusions ends in an acute reaction, with as many as 1 in 600,000 ending in death. Delayed reactions occur in as many as 1 in

5000 transfusions but are less often fatal.

Traditional blood typing has other limitations. Some people have a weak or partial version of the D antigen that is difficult to detect. Misdiagnosing D+ donor blood as D- could cause serious problems. A D- person given such mislabeled blood will make antibodies that would react to future transfusions of D+ blood. Furthermore, it's important for women to accurately know their RhD status: If a D- woman had a D+ baby, any subsequent D+ fetus is at risk because the mother will have made antibodies to the D antigen during the first pregnancy. Because serology has trouble deciphering RhD status, says Avent, as many as 40% of pregnant women may receive unnecessary drug treatment, which can have side effects, to limit their antibody production.

Genotyping offers an alternative. By sequencing the 31 genes that determine the surface antigens, hematologists can classify the blood into the 29 groups. For example, a single gene on chromosome 9 controls the ABO antigens, and the versions inherited from each parent determine a person's type. Other blood groups are determined by a single nucleotide polymorphism (SNP), which changes just one base in an antigen-determining gene. Insertions or deletions of DNA within other blood-group genes lead to different antigens. By identifying these different versions of a gene, or alleles, genotyping can predict a person's blood types in fine detail.

Advocates of genotyping predict that with better matched blood, they can prevent between 80% and 90% of alloimmunization and also eliminate partial and weak versions of the D antigen from the blood supply. Rosse says such tests might have helped his patient. "If we had his genotype, we could have matched it with a donor's genotype," Rosse says. "We might have found a match."

Europe takes the lead

Researchers have been using polymerase chain reaction (PCR)-based techniques to genotype blood samples in the lab for a decade. But if blood genotyping is extended into widespread clinical use, cheaper and faster methods will be needed.

In 2002, the European Union gave €2.35 million to BloodGen, a consortium of universities and blood centers across Europe that planned to standardize blood genotyping techniques and prove that they beat serology. The consortium, led by Avent, has since developed the BLOODchip, a gene chip that tests a person's DNA using blood samples. PCR alone would require 60 or more tests to determine the blood type as comprehensively

as the chip can with just one test. The current chip, produced by the company Progenika, looks at nine blood groups, including the genes that code for the A, B, and D antigens.

In initial tests, the BloodGen team genotyped 1000 blood samples and found 42 cases that conflicted with serology, some of which were in the Rhesus blood group. Further analysis revealed that two of these errors were the fault of genotyping and 40 were the fault of serology. Progenika has full clinical approval for seven of the blood groups on its chip and is currently seeking approval for RhD. The next round of tests will look at 3000 samples.

BioArray Solutions in Warren, New Jersey, has developed another genotyping product called BeadChip that tests for 11 blood groups but not A, B, and D. Although the omission is in part practical, some researchers have said it reflects a larger debate in the scientific community about the potential of genotyping to replace, as opposed to supplement, serology.

A bloody debate

Avent is adamant that genotyping will replace serology for most blood groups within the decade and in some cases sooner. He predicts European hospitals will demand blood genotyping of patients expecting to receive a transfusion. "For multitransfused patients, I expect a change in policy this year," Avent says. "Then I would expect other vulnerable groups [such as pregnant women to be] tested. Then I would like to see blood centers genotyping cohorts of donors."

Researchers in the United States are taking a more cautious approach. They believe genotyping has a way to go for the A, B, and D antigens. Researchers have identified more than 100 alleles for the ABO blood type and more than 200 for the Rhesus system, and new mutations are discovered frequently, which makes some people nervous about relying on current gene chips. "I don't think we have found all the alleles," says Marion Reid, an immunohematologist at the New York Blood Center.

Avent predicts that within the next few years, researchers will have compiled the majority of those alleles, and he adds that

genotyping is already superior for identifying the D antigen. Willy Flegel, a transfusion medicine specialist at University Hospital Ulm in Germany, agrees. He has genotyped more than 46,000 blood donations identified as D- based on serology and found that 47 are actually D+. "The smart serologist will apply molecular techniques now for the benefit of the patient," Flegel says. "The scientific arguments are clearly in favor of genotyping."

Connie Westhoff, scientific director at the American Red Cross, concurs that genotyping is the future, but she doesn't expect that the United States will fully adopt it for another 15 years at least, and even then, she does not expect the country to completely abandon serology. "We would never

throw out our old toolbox," she says.

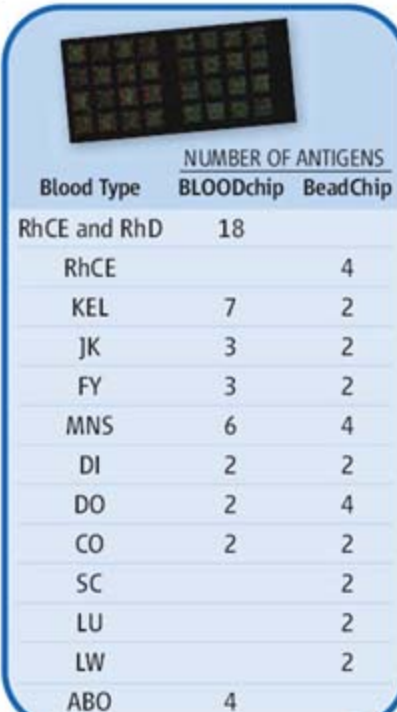
Moving forward

A few countries are already forging ahead with blood genotyping on a large scale. Canada is among the leaders. In December, for example, Quebec announced plans to genotype 22,000 blood donors. Currently, when a patient with a rare blood group needs a transfusion, practitioners have to blindly order blood from banks based on ABO and D grouping alone and then test it with more detailed serology on site. This trial-and-error approach wastes time and money. The screening should make the process faster by narrowing initial blood selection, but serological testing will still serve as a follow-up.

Jean-Pierre Cartron, scientific director at France's national institute of blood transfusion, says his blood center will be genotyping using both the BLOODchip and the BeadChip by the beginning of April. "We want to see them together on the same population of donors and patients," Cartron says.

Flegel says the ABO group will take time to work out, but for all other groups, genotyping should begin now. Only by using the available genotyping tools can hematologists locate unknown alleles and make future gene chips more accurate. "We need to learn what we are missing," he says. He agrees that hematologists need to proceed with caution. But, he says, "we don't need to wait."

—ELIZABETH QUILL



Blood Type	NUMBER OF ANTIGENS	
	BLOODchip	BeadChip
RhCE and RhD	18	
RhCE		4
KEL	7	2
JK	3	2
FY	3	2
MNS	6	4
DI	2	2
DO	2	4
CO	2	2
SC		2
LU		2
LW		2
ABO	4	

Chipping away. Two new gene chips (BLOODchip above) can test for many blood groups beyond ABO.

COMPUTER SCIENCE

Cryptologists Cook Up Some Hash for New 'Bake-Off'

A worldwide competition aims to keep the algorithms for authenticating electronic documents a jump ahead of forgers' ability to defeat them

In November, three Dutch cryptologists published on their Web site a "digital fingerprint" of their prediction of the winner of the United States's 2008 presidential election. According to them, the next commander in chief will be 3D515DEAD7AA16560ABA3E9DF05CBC80.

After the election, these modern-day Nostradamuses will use a simple mathematical procedure called a hash function, which is available on nearly all computers, to show that a PDF document created before the election has that hexadecimal number as its digital fingerprint. This will prove that they knew the winner all along, because the PDF document could not have been altered after the election. Any such change would make its fingerprint, known as a hash value, no longer match the one that has been published.

But it's all a trick. The "Nostradamus attack" by Marc Stevens, Arjen Lenstra, and Benne de Weger is designed to highlight a serious problem in cryptography: The so-called hash functions that many of the world's computers use for authenticating documents are dangerously out of date. The cryptologists prepared 12 separate documents, one saying that John McCain will win, one naming Hillary Clinton, and one even predicting Paris Hilton. By carefully tweaking the contents of each PDF document, in a way not obviously noticeable, they made it so each one generates the same digital fingerprint, computed by a widely used hash function called MD5. As the Dutch effort shows, the ability to produce multiple documents with the same fingerprint renders MD5 useless for authentication.

Without the ability to authenticate files, such as passwords and online transactions, Internet commerce would be seriously threatened. Therefore, the U.S. National Institute of Standards and Technology (NIST) in November announced a worldwide competition to select a new standard for hash functions, which is expected to conclude by 2012. The winner will be certified for U.S. government use, and if past history is any guide, that will make it a de facto standard for the rest of the world.

NIST held a similar "bake-off" (as some cryptologists called it) from 1997 to 2000 to select a new standard cipher for government use, called the Advanced Encryption Standard (AES). "The AES competition was the most fun I've ever had in cryptography," says Bruce Schneier of BT Counterpane in Santa Clara, California, who designed one of the five AES finalists and who plans to enter the hash-function competition as well. "Think of it as a giant cryptographic demolition derby: A bunch of us put our best work into the ring, and then we beat on each other until there was only one standing. ... I personally learned an enormous amount about [cipher design] from the AES competition, and we as a community benefited immeasurably."

Like the previous competition, the new bake-off offers no financial reward, and the submissions must be unpatented so that they can be incorporated into any software. The real payoff to the winner will be prestige. The AES competition drew 15 entries, and the hash-function competition is expected to attract as many as 50.

Although the contests will be similar, the products could hardly be more different. A cipher, like AES, encrypts data in such a way that it can be recovered but only by someone with a key. Hash functions, on the other hand, are not meant to be reversible, and they have no secret key. They merely show that a document is what it claims to be. They tend to be much simpler than ciphers and have a much broader range of applications—so many that they are often called the "duct tape" or "Swiss army knife" of cryptology.

Hash functions work by converting any string or message of 1s and 0s to a new string, usually much shorter. For example, the function might take a gigabyte MPEG movie file and boil it down to anywhere from 128 to 512 bits. In the current government standard, called SHA-1, the output, known as the hash, hash value, or hash sum, is 160 bits, whereas for MD5 it is

128 bits. The hash should look more or less random so that no one can guess what the original message said. On the other hand, it should be generated in a completely deterministic manner so that anybody who knows the hash function used can have a computer verify that the hash value matches the hashed file's contents. In addition, it should be staggeringly unlikely that any other document hashes to the same value. When that happens, cryptographers call it a "collision."

Compressing gigabyte-sized files to 160 bits inevitably results in many collisions. But the key point is that it is virtually impossible to find one by random search. That is because there are so many possible hash values— 2^{160} of them for a 160-bit hash sum. For such a hash value, the number of movies you would have to hash before generating the same hash twice would be roughly 2^{80} , or a trillion a year for a billion years. Not even Bollywood is that prolific.

However, a hash function can be defeated if it is possible to deliberately create two colliding documents. If this takes more than 2^{80} attempts, then the hash function is considered secure, because such an attack is no better than random guessing. However, if a collision can be found in less than 2^{64} tries, then a supercomputer, or a very large network of personal computers, might be able to do it in a year. If it takes less than 2^{32} tries, one can do it on a PlayStation in a few minutes—as the Dutch "Nostradamuses" did.

The current standard, SHA-1, borrows its basic architecture from MD5, which was invented in 1992. You feed your message into a device that compresses and randomizes 512 bits at a time. You add another piece of your file to the first piece and feed it through again, and you repeat this procedure until you run out of message.

Iterative algorithms, like MD5 and SHA-1, are very easy to program and quick to run. But recently, they have come under heavy attack from cryptologists. In 2004 and 2005, cryptologist Xiaoyun Wang of Shandong University in China showed that MD5 could be cracked in fewer than 2^{40} steps, and SHA-1 in fewer than 2^{64} . No one has produced an actual collision in SHA-1, but a search, using the spare time of many personal computers, is under way at Graz University of Technology in Austria.

"If you find a [SHA-1] collision, people in industry will be forced to upgrade their products," says Bart Preneel, a cryptologist at

"It may turn out that they aren't broken or can't be broken, but we didn't want to get caught out on the wrong side."

—WILLIAM BURR, NIST

Hash of the Future?

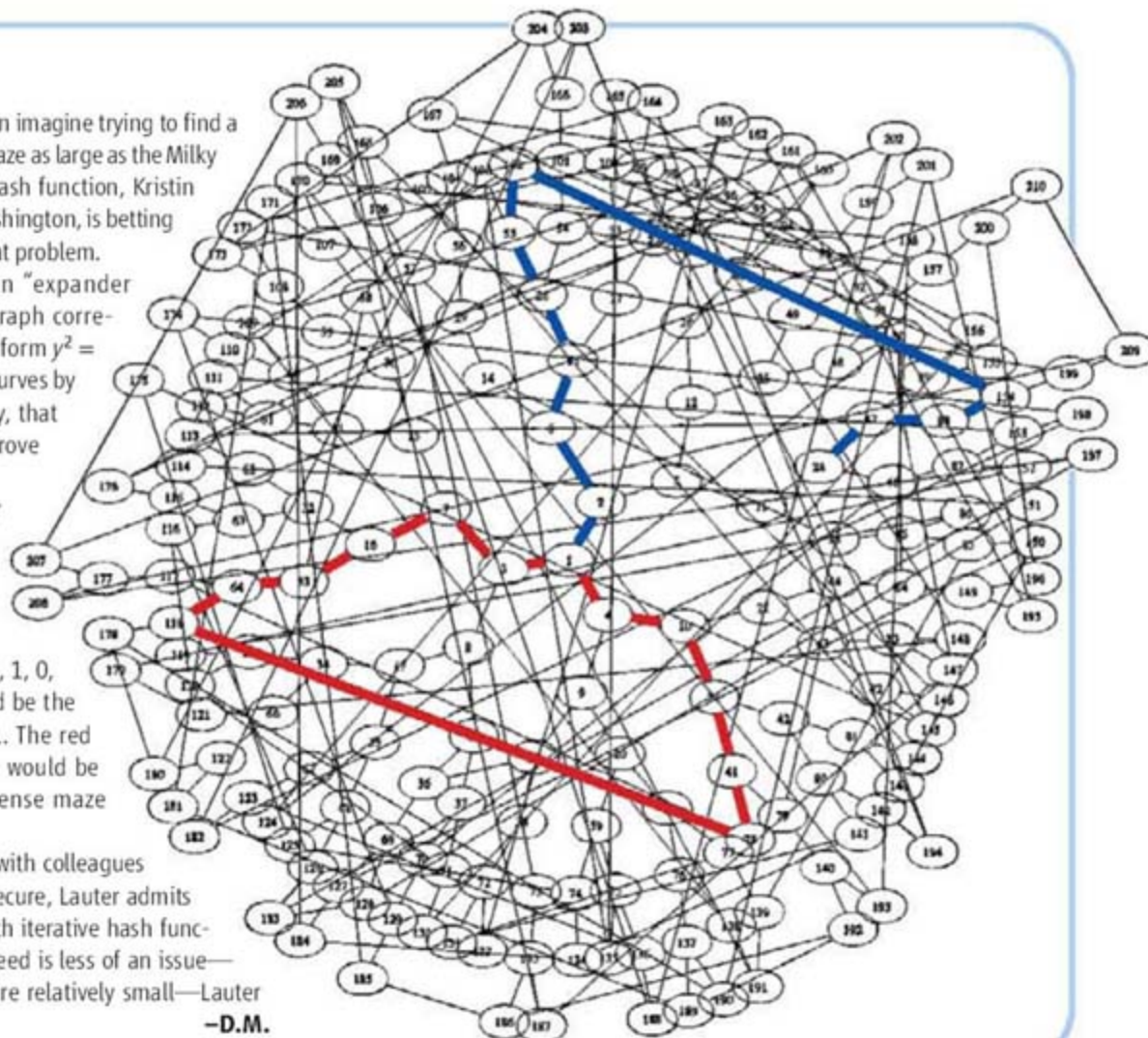
Have you ever struggled to solve a maze? Then imagine trying to find a path through a tangled, three-dimensional maze as large as the Milky Way. By incorporating such a maze into a hash function, Kristin Lauter of Microsoft Research in Redmond, Washington, is betting that neither you nor anyone else will solve that problem.

Technically, Lauter's maze is called an "expander graph" (see figure, right). Nodes in the graph correspond to elliptic curves, or equations of the form $y^2 = x^3 + ax + b$. Each curve leads to three other curves by a mathematical relation, now called isogeny, that Pierre de Fermat discovered while trying to prove his famous Last Theorem.

To hash a digital file using an expander graph, you would convert the bits of data into directions: 0 would mean "turn right," 1 would mean "turn left." In the maze illustrated here, after the initial step 1-2, the blue path encodes the directions 1, 0, 1, 1, 0, 0, 0, 0, 1, ending at point 24, which would be the digital signature of the string 101100001. The red loop shows a collision of two paths, which would be practically impossible to find in the immense maze envisioned by Lauter.

Although her hash function (developed with colleagues Denis Charles and Eyal Goren) is provably secure, Lauter admits that it is not yet fast enough to compete with iterative hash functions. However, for applications in which speed is less of an issue—for example, where the files to be hashed are relatively small—Lauter believes it might be a winner.

—D.M.



Katholieke Universiteit Leuven in Belgium. Anticipating such a breakdown, Microsoft in 2005 banned both SHA-1 and MD5 from new products and has removed MD5 from all its current products, says Kristin Lauter, head of the Cryptography Group at Microsoft Research in Redmond, Washington. Fortunately, a good backup is already available. In 2004, NIST issued several new standards, collectively called SHA-2, which are more secure than SHA-1 because they produce longer hashes (up to 512 bits instead of 160).

But NIST worries that SHA-2 could eventually fall, too. "Everything that has been attacked is in the same family," says William Burr of NIST's Security Technology Group. "It may turn out that they aren't broken or can't be broken, but we didn't want to get caught out on the wrong side."

After extensive debate, including two international workshops in 2005 and 2006, NIST decided that a new competition could turn up completely new approaches to hash functions. "We'll be reluctant to pick something that looks just like SHA-2," says Burr. "We want some biodiversity."

Although no designs have been formally

submitted yet—the deadline is in October—experts predict that most entrants will continue to be iterative algorithms subtly retooled to defeat the new kinds of attacks. For instance, Preneel's RIPEMD—one of the few first-generation hash functions still standing—performs two parallel iterations, making it difficult for an attacker to figure out which one to attack.

A second approach, called "provably secure" hash functions, derives its presumptive security from math problems that are considered to be hard to crack (see sidebar, above). This type of algorithm typically does not require multiple iterations, but it does require cryptologists to put their faith in a mathematical "black box." Also, such algorithms tend to be slower than iterative algorithms because they require a more elaborate calculation—even though it is performed only once. Speed is at a premium for hash functions, as they are typically used to tag a document in the split-second it's electronically transmitted.

Not surprisingly, mathematicians love provably secure systems, whereas cryptologists have little use for them. "They are typi-

cally only provable with respect to one property but are weak with respect to other properties," says Joan Daemen of STMicroelectronics, co-winner of the AES competition. For instance, a "provably secure" hash developed by Lenstra and his colleagues, called Very Smooth Hash (VSH), was compromised last year when Markku-Juhani Saarinen at a Spanish company called Kinamik showed that it was easy to find "near-collisions" in VSH. In practice, engineers often truncate a long hash value to a shorter one, assuming that the truncated hash will inherit the long one's security. Saarinen's result means that they can't count on that with VSH.

In the final analysis, what makes it so hard to come up with good hash functions—and prove they work—is that they are expected to do so many things. "You expect them to do everything and blame them when they don't work," says Preneel. Perhaps a 4-year bake-off will be just what the chef ordered to make some new hash that will satisfy everybody's tastes.

—DANA MACKENZIE

Dana Mackenzie is a freelance writer in Santa Cruz, California.

Blackburn biography

1486



Primate exploitation

1487



How humans learn

1493



LETTERS | BOOKS | POLICY FORUM | EDUCATION FORUM | PERSPECTIVES

LETTERS

edited by Jennifer Sills

The Right Pitch for Saxophonists

THE BREVIA BY J. M. CHEN *ET AL.* ("EXPERIENCED SAXOPHONISTS learn to tune their vocal tracts," 8 February, p. 776) substantially adds to the debate, now more than 25 years old, on the influence of the vocal tract on the tone production of single-reed instruments. I thought it might be useful to offer some practical experiences on the subject as a jazz saxophonist who has been playing professionally for nearly 50 years. Professional saxophonists have always been looking for a scientific explanation of the phenomenon often described by myself and others as "singing through the horn." This is a totally instinctive and rather colorful expression that describes rather well the use of the vocal tract in the art of good tone production on the saxophone or other reed instruments.

I first became aware of the importance of using my vocal tract while playing in the saxophone section of a famous British big band in the early '60s. One night, after some thought on the subject, I discovered that if I mentally "sang" the notes, I was able to play much better in tune and with a fuller and more centered tone. It made sense that by almost silently singing the notes, the sound would be more focused. Of course, for many years now, the importance of this aspect of technique has been well known, but it has not been well understood.

I have always taught saxophone pupils to "sing through their horns" but now elaborate with more precise instructions. In general, the higher the note on the instrument, the more one has to lift the back of the tongue and restrict the size of the oral cavity. This explains why

using the vocal tract is especially important when playing in the "altissimo" register of the saxophone.

The saxophone is without doubt the most "flexible" of all the reed family; it has the widest range of possible tone colors of all wind instruments. This flexibility is the reason why the greatest jazz saxophonists have such uniquely personal timbres. One only needs to hear "masters" like John Coltrane, as well as Coleman Hawkins or Stan Getz, to realize the saxophone can produce sounds of enormous variety; sounds whose individuality comes closer to the human voice than on any other instrument.

I have always maintained that the saxophone is one of the easiest instruments to pick up and get a tune out of, but the hardest of all to play with a truly great and personal sound. Much of the art of saxophone playing relies on an impeccable understanding and control of the vocal tract and the ability to almost literally "sing through the horn."



Singing through the horn. Peter King plays his saxophone.

PETER KING

15 Oakhill Place, London SW15 2QN, UK. E-mail: pj2king5te@aol.com

Second Basket's Negative Impact

ON 1 JANUARY 2008, THE GERMAN "SECOND Act Governing Copyright in the Information Society" ("Second Basket") took effect. Second Basket limits the circulation of scientific documents and will substantially reduce scientific communication among researchers. For example, libraries or other document providers may send copies electronically (e.g., in PDF format) only if the publisher of a particular work does not itself offer the work online in a clear manner and on reasonable terms. In principle, the Copyright Act strengthens the protection of authors' intellectual rights; however, most scientists do not want this protection if it

comes at the expense of efficient and effective circulation of scientific information. Effective science relies on effective communication, often quantified by the term "impact" (1).

Who is really protected by limiting the circulation of scientific knowledge? Publishers benefit most. The Copyright Act requires them to deliver scientific work on "reasonable terms," making possible a wide spectrum of business terms and conditions. Transfer agreements allow publishers to claim the copyright. A paradoxical situation emerges: National research foundations pay scientists to write papers, but the scientists must transfer the copyrights and then buy back their own papers in order to provide access to them. The international

community should create a suitable framework for international copyrights in a networked world (2).

CLAUS-CHRISTIAN CARBON

Department of Psychology, University of Vienna, Vienna, Austria, and TU Delft, Delft, Netherlands.

References

1. E. Garfield, *Science* **178**, 471 (1972).
2. M. Seadle, *Library Hi Tech* **25**, 298 (2007).

Going Public with the Scientific Process

THE IDEA OF USING FRAMING STRATEGIES TO communicate science to the public has recently been taken up in scientific forums

CORRECTIONS AND CLARIFICATIONS

Cover Caption: (1 February). The cover image was originally published as the right side of figure 8 from *Chem. Mater.* **19**, 5485 (2007), copyright 2007 American Chemical Society.

Perspectives: "From molecules to memory in the cerebellum" by D. J. Linden (19 September 2003, p. 1682). As a result of an editorial oversight, a credit for the second figure was missing. The figure was adapted from figure 2 of D. A. McCormick et al., *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2731 (1982).

(1, 2), the mainstream media (3), and the blogosphere (4, 5). Most participants in the framing science debate limit their notion of scientific information to scientific facts. However, confining science messages to just the facts interferes with public understanding of science as a systematic, logical process of human inquiry and effaces the distinction between data and scientists' reasoning about data. To communicate successfully, we should focus on scientific process by emphasizing two important elements of scientific rationality: skepticism and dynamicism (6, 7).

Scientists deliberately integrate skepticism into their procedures by trying to refute their own hypotheses, retaining them only when confronted with compelling evidence sought through carefully controlled procedures. Scientists tend to shy away from revealing the intrinsic skepticism of science to the public, fearful that it will open the door to doubt about the validity of their conclusions. But communicating only the facts of science (framed or unframed) destabilizes public confidence in science. A fact doesn't allow science communicators to reveal, justify, and ultimately promote the skeptical reasoning process that helps make scientists more confident that their reasoning is correct.

Science is also dynamic; it is a cumulative enterprise that requires scientists to situate their instrumental activities and interpretations against the evidence that has come before and to alter them in light of new evidence. Insisting that new data be interpreted within the context of past and future data will ferret out and correct error over time, but it means that a fact cannot, by definition, be anything more than the (ephemeral and fallible) consensus of scientists at a given point in time. A "just the facts" strategy can and often does backfire, ultimately fueling public alienation from science. When scientists inform the public of "facts" (like the "fact" widely disseminated in the 1970s that all dietary fats are bad for us), and then that "fact" is refined or altered (now we're told olive oil is good for us), the public is justifiably confused. Studies suggest that the public tends to regard normal scientific refinement and self-correction as equivocation or incompetence (8–10). Instead of sweeping

uncertainty under the rug, science communicators should help the public understand the logical and systematic procedures by which scientists confront it.

The true majesty and promise of science lies in its systematic, logical, skeptical, and dynamic reasoning procedures. "Successful" science communication should not be regarded as any message that enlists public support for science. Rather, we should define "success" in scientific communication as achieving a public that celebrates scientific reasoning procedures.

RUTH CRONJE

Scientific and Technical Writing Program, University of Wisconsin–Eau Claire, Eau Claire, WI 54701, USA.

References

1. M. C. Nisbet, C. Mooney, *Science* **316**, 56 (2007).
2. M. C. Nisbet, D. Scheufele, *The Scientist* **38**, 38 (2007).
3. M. C. Nisbet, C. Mooney, "Thanks for the facts. Now sell them." *The Washington Post*, 15 April 2007; Outlook section, p. B03.
4. AAAS News Blog, "Science has a 'serious marketing problem,' says Google founder Larry Page," 17 February 2007; available online at www.aaas.org/news/releases/2007_ann_mtg/127.shtml.
5. P. Z. Myers, "What if the right role for science is to shatter the frame?" *Pharyngula*, 7 April 2007; http://scienceblogs.com/pharyngula/2007/04/what_if_the_right_role_for_sci.php.
6. J. Habermas, *The Theory of Communicative Action, Vol. 1: Reason and the Rationalization of Society*, T. McCarthy, transl. (Beacon Press, Boston, 1984).
7. H. I. Brown, *Rationality* (Routledge, New York, 1988).
8. B. Wynne, *Environment* **31**, 10 (1989).
9. B. B. Johnson, P. Slovic, *Risk Anal.* **15**, 485 (1995).
10. B. B. Johnson, *Risk Anal.* **28**, 781 (2003).

Hoyle's Role in B²FH

IN JULY 2007, THE NUCLEAR ASTROPHYSICS 1957–2007 conference (1) commemorated the 50th anniversary of the publication of "Synthesis of the elements in stars," by Burbidge, Burbidge, Fowler, and Hoyle (referred to by the shorthand B²FH) (2). In response, D. D. Clayton ("Hoyle's equation," Perspectives, 21 December 2007, p. 1876) wrote that a key paper by Hoyle (3) in the development of the theory of stellar nucleosynthesis has been undercited and, by implication, that not enough credit has been given to Fred Hoyle. I agree.

As one of the only two survivors of B²FH, I would like to provide some additional comments. First, the theory of stellar nucleosyn-

thesis is attributable to Fred Hoyle alone, as shown by his papers in 1946 (4) and 1954 (3) and the collaborative work of B²FH (2). In writing up B²FH, all of us incorporated the earlier work of Hoyle (3, 4).

In my view, Hoyle's work has been undercited in part because it was published in an astrophysical journal (3), and a new one at that (the very first volume, in fact), whereas B²FH was published in a well-established physics journal, *Review of Modern Physics*. When B²FH was first written, preprints were widely distributed to the nuclear physics community. Willy Fowler was very well known as a leader in that community, and the California Institute of Technology already had a news bureau that knew how to spread the word.

As I pointed out at the meeting in July (5), Margaret Burbidge and I wrote the first draft of B²FH. We deliberately incorporated extensive observations and experimental data supporting the theory, and Hoyle and Fowler worked extensively on the early draft to see that all of the work was covered. There was no leader in the group. We all made substantial contributions, and Hoyle was entirely happy with the result.

Hoyle should have been awarded a Nobel Prize for this and other work. On the basis of my private correspondence, I believe that a major reason for his exclusion was that W. A. Fowler was believed to be the leader of the group. As I stated in Pasadena, this was not the case (5).

GEOFFREY BURBIDGE

Center for Astrophysics and Space Sciences, University of California, San Diego, La Jolla, CA 92093–0424, USA.

References

1. Nuclear Astrophysics 1957–2007, California Institute of Technology, Pasadena, CA, 23 to 27 July 2007 (www.na2007.caltech.edu).
2. E. M. Burbidge, G. Burbidge, W. A. Fowler, F. Hoyle, *Rev. Modern Phys.* **29**, 547 (1957).
3. F. Hoyle, *Astrophys. J. Suppl.* **1**, 121 (1954).
4. F. Hoyle, *Mon. Not. R. Astron. Soc.* **106**, 343 (1946).
5. G. Burbidge, 2008, paper presented at the Nuclear Astrophysics 1957–2007 conference, California Institute of Technology, Pasadena, CA, 23 to 27 July 2007; *Pub. Astron. Soc. Aust.*, in press.

Letters to the Editor

Letters (~300 words) discuss material published in *Science* in the previous 3 months or issues of general interest. They can be submitted through the Web (www.submit2science.org) or by regular mail (1200 New York Ave., NW, Washington, DC 20005, USA). Letters are not acknowledged upon receipt, nor are authors generally consulted before publication. Whether published in full or in part, letters are subject to editing for clarity and space.

COMMUNICATING SCIENCE

Fresh Renderings of Physics

David Kaiser

During its first decade in print, Stephen Hawking's *A Brief History of Time* (1) sold an astonishing 12 million copies in English-language editions alone and millions more in a host of other languages. Once upon a time, those were jaw-dropping sales figures (at least before Harry Potter). Scholars have often pointed to the runaway success of Hawking's book to explain a boom in popular-science book publishing during the 1990s, when more than two dozen nonfiction books on scientific topics became bestsellers in both the United Kingdom and the United States. Yet, as literary scholar Elizabeth Leane (the University of Tasmania) rightly notes in her tightly argued and fascinating study, *Reading Popular Physics*, popular-science

and especially the popular books about physics on which she focuses—played major roles in the overheated “science wars” of the 1990s, which pitted ideas from a variety of humanistic disciplines against a traditional view of how scientific knowledge is generated, vetted, and stabilized.

The book opens with three chapters placing the popular-science boom in various contexts, including the science wars. Exhibit A: the famous (or infamous) “Sokal affair” of 1996, in which New York University physicist Alan Sokal composed an essay spoofing recent trends in the cultural studies of science and managed to get it published in an (unrefereed) cultural studies journal. Leane, who completed a degree in theoretical physics before shifting to literary studies, navigates the thickets of the conflicts with patience and intelligence, demonstrating clearly that exponents of both sides frequently resorted to superficial caricatures of each other's positions. Moreover, as she reminds us, many of the first battles played out in popular-science venues.

The second half of the book shifts to close textual studies. Popularizations often serve as bridges for cross-disciplinary borrowings. Yet they are hardly transparent representations of underlying scientific research. They are more like fresh renderings, akin to translating a poem into a different language (2). Thus, Leane argues, it is crucial to pay particular attention to the literary strategies used by popular-science authors.

Leane's first theme concerns the use of metaphor and rampant anthropomorphism in popularizations of quantum theory. Sometimes the metaphors can be as subtle as an author's choice of verb, such as the common descriptions of an electron “knowing” when to jump from an excited orbit to a lower one in the Bohr model of the atom or a photon “choosing” which slit to pass through in a double-slit experiment. The bulk of this chapter concerns Gary Zukav's well-known *The Dancing Wu Li Masters* (3)—one of the few to which literary scholars continue to turn when

looking for an accessible introduction to the topic—which dips more than most in the anthropomorphic waters.

Next comes a study of narrative structure and the use of mythic tropes in popular treatments of big bang cosmology. Steven Weinberg's *The First Three Minutes* (4) follows the classic narrative style of tragedy: All threads are tied up in the end, leading to Weinberg's famously pessimistic conclusion. Hawking's *Brief History of Time*, meanwhile, unfolds like a comedy: Seeds are planted

early for a happy union at the end—in this case not a marriage of protagonists, but a “theory of everything.”

Leane's final chapter focuses on popular authors' styles of depicting characters. She produces lovely parallels between the tone and even specific word choices and colloquialisms in James Gleick's *Chaos* (5), M. Mitchell Waldrop's *Complexity* (6), and the hard-boiled noir detective fiction of Dashiell Hammett and Raymond Chandler. Like the gumshoe private investigators, the characters in the popularizations are depicted as working alone, outside society, and immune to its superficialities and cultural codes of conduct. As Leane points out, this choice of stencil carries consequences. It leads to strongly gendered male stories, with virtually no women in view. And it stakes a claim about scientists' proper relation to society—namely, they should remain outside it and beyond its defiling reach—just at the moment that the science wars pushed the question of science and society to center stage.

The book's lengthy bibliography offers further riches for specialists, while the main text remains accessible to readers with no background in literary criticism. *Reading Popular Physics* is a sophisticated, engaging book on the ways we talk about science—and their consequences.

References and Notes

1. S. Hawking, *A Brief History of Time: From the Big Bang to Black Holes* (Bantam, London, 1988).
2. As Leane notes, this metaphor is from Jon Turney, *Public Understand. Sci.* 13, 331 (2004).
3. G. Zukav, *The Dancing Wu Li Masters: An Overview of the New Physics* (Morrow, New York, 1979).
4. S. Weinberg, *The First Three Minutes: A Modern View of the Origin of the Universe* (Basic Books, New York, 1977).
5. J. Gleick, *Chaos: Making a New Science* (Viking, New York, 1987).
6. M. M. Waldrop, *Complexity: The Emerging Science at the Edge of Order and Chaos* (Simon and Schuster, New York, 1992).



publishing had been flourishing for at least a decade before Hawking's book appeared. Hawking rode a wave—and rode it with more gusto than most could have imagined—but he did not invent the wave itself.

The focus of *Reading Popular Physics* is the liminal status of best-selling science books, caught betwixt and between recognized genres: not exactly original contributions to scientific research, but rarely mistaken for high-brow literature, either. Instead they are hybrids, borrowing certain conventions from both sides of what C. P. Snow famously called the “two cultures.” As such, Leane demonstrates, popular-science books—

The reviewer is at the Program in Science, Technology, and Society and the Department of Physics, Building E51-185, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA. E-mail: dikaiser@mit.edu

MOLECULAR BIOLOGY

On Ends and Means

Michael A. Goldman

Elizabeth H. Blackburn, telomere, and telomerase are not yet household names, but odds are we'll see Blackburn in Stockholm before we see federally funded research on embryonic stem cells. Blackburn's groundbreaking work in telomere biology is a remarkable story worth telling. Beyond this, Catherine Brady's *Elizabeth Blackburn and the Story of Telomeres* offers a commanding account of an inspiring effort to overcome gender bias along with advice about doing science, conquering academic politics, and taking responsible positions on science policy.

"For her," says Brady, "a feeling of awe for the mystery of life, symbolically embedded in the 'secret code' of the genes, was wedded to curiosity and a methodical effort to dismantle that mystery." Fred Sanger invited Blackburn to pursue graduate study in his laboratory at the Medical Research Council's Laboratory of Molecular Biology in Cambridge. There she rubbed shoulders with the likes of Sydney Brenner, James Watson, and Francis Crick; got into DNA sequencing at ground zero; and met her future husband, John Sedat. You get the impression that she was intelligent and, although mostly silent, able to hold her own without difficulty in a British tradition that invited the students to tea and lunch with the eminent scientists of the day. She no doubt benefited from picking up information even if she wasn't always a lively participant.

As a postdoc in Joe Gall's lab at Yale, Blackburn applied her knowledge to determining the nucleotide sequence of telomeres in *Tetrahymena*. She went on to dominate the field, sequencing telomeres in several other organisms and, with her graduate student Carol Greider, identifying the telomerase enzyme.

Brady pays close, but not unwarranted, attention to the dynamics of a woman navigating a male-dominated science. She paints a picture of Blackburn steering clear of the most competitive aspects of molecular biology,



Elizabeth Blackburn and the Story of Telomeres
Deciphering the Ends of DNA

by Catherine Brady

MIT Press, Cambridge, MA,
2007. 412 pp. \$29.95, £19.95.
ISBN 9780262026222.

carving out a niche of her own, and then becoming a titan in that field. Brady is fervent in interpreting every event in Blackburn's life as a conquest of her disadvantages as a woman rather than as a natural evolution of her excellence as a scientist: "That Blackburn, so eminent in her field, still could not see herself as powerful speaks volumes about the difficulties that continue to beset women as they enter the higher ranks of academia. She couldn't imagine a negotiating strategy other than threatening to leave...." Sometimes, no matter who you are, there is no other strategy.

Although Brady (a professor in the writing program at the University of San Francisco) has done an admirable job of educating herself in the science, readers desiring a good understanding of telomeres and telomerase will need some background—say, a couple of semesters of college biology. Yet all should come away comprehending the main scientific points, recognizing the potential clinical applications, and appreciating how science is done and just how painstaking work at the laboratory bench is.

What most researchers will value most from the book is its compilation of a master's wisdom about doing science. Hints and quotable lines abound, from how to think about data to how to navigate a career. Brady explains that Blackburn "tends to 'roam' the data," "exploring a single experimental question from every angle and ... synthesizing a wide range of potentially relevant information." Blackburn captures a key difference between the biological and physical sciences when she notes "You can't operate from first principles, as physicists can, but have to test

logic against precedent and the data itself. Proof in the purest sense is something that comes in a messy way."

Blackburn faults the granting and publishing establishments in which she has been quite successful. She notes that the National Institutes of Health are risk-averse, but "countering this tendency ... is the fact that the grant itself is not a contract." Reviewers sometimes abuse the process. "Who knows what knives are out there, under the cloak of anonymity?" With disdain Blackburn remarks, "I don't even really know what the impact factor is.... It's the bottom-feeders who talk about the impact factor."

A staid member of the President's Council on Bioethics, Blackburn holds that "[i]t is important for scientists to participate in policy debates because they can evaluate the data." Her insistence on accuracy put her at odds with council chair Leon Kass, whose "yuck factor" replaced reason when presidential policy on embryonic stem cell research conflicted with science. The result was her dismissal from the Council and another embarrassment for the president.

Brady never shows us the dark side of Blackburn, but the encounters I've had with Blackburn's former fellows, students, and colleagues suggest that there isn't one. She always supported her students and fellows whether they chose to follow her career path or not. She thinks that it is important to have competent scientists in every walk of life. Leading a new field and training the people who will take it into the future is a tremendous contribution. But Blackburn's academic legacy is notable because it has begun to spread far beyond academe. I'm waiting for a sequel.

The reviewer is in the Department of Biology, San Francisco State University, San Francisco, CA 94132-1722, USA. E-mail: goldman@sfsu.edu

10.1126/science.1154779

CREDIT: ELIZABETH FALL

SCIENCE PRIORITIES

Inappropriate Use and Portrayal of Chimpanzees

S. R. Ross,^{1,2} K. E. Lukas,³ E. V. Lonsdorf,^{1,4} T. S. Stoinski,^{5,6} B. Hare,⁷ R. Shumaker,^{8,9} J. Goodall¹⁰

In North America alone, about 2300 chimpanzees live in a variety of settings from accredited zoological parks to laboratories and sanctuaries. However, in 44 of the 50 states in America, chimpanzees can also be privately owned as pets and/or used as actors and photographer's props in the entertainment and media industry (1). In movies, television shows, and advertisements, chimpanzees are often depicted as caricatures of humans, dressed in clothes and/or photographed in contrived poses. For example, chimpanzees are portrayed as misbehaving business executives in the popular "Careerbuilder" advertisements. More recently, chimpanzees were shown dressed in hats while reading an issue of *Science* magazine in a promotional campaign by the American Association for the Advancement of Science (AAAS) (although, it should be noted that the campaign was halted when AAAS was made aware of objections). Such inappropriate portrayals are viewed by millions of people annually and may influence the way in which members of the general public perceive this endangered great ape.

In 2005, a survey (see the table, left side) was conducted at the Regenstein Center for African Apes (RCAA) at the Lincoln Park Zoo (Chicago, IL). The full survey was made up of 56 questions with the intent of assessing the effect of a new facility on visitor knowledge and attitudes toward apes (2). The final question of the survey asked respondents to select which of three great ape species (chimpanzees, gorillas, and orangutans) were considered endangered in the wild. Labeled photographs were used to ensure that species identification

was correct, and the order of the species list was randomized. Of those choices, 95% of respondents thought gorillas were endangered, 91% thought orangutans were endangered, but only 66% believed chimpanzees to be endangered. This species-level difference was significant by a two-way χ^2 analysis for chimpanzees versus gorillas (CvG), $\chi = 37.726$, $df = 1$, $P < 0.001$; for chimpanzees versus orangutans (CvO), $\chi = 22.588$, $df = 1$, $P < 0.001$ (see table, right) (2). A follow-up question addressed the potential explanation for such a difference. Respondents were

informed that, in fact, all three great apes were classified as endangered and then asked for a reason why they thought a particular ape was not considered in this category. No prompting with answers was provided, and all responses were recorded by the interviewer. Of the 250 respondents who were willing to provide explanations for their choice, the most common reason for the category chosen (35%) was that chimpanzees were commonly seen on television, advertisements, and movies and, therefore, must not be in jeopardy.

The results were later duplicated in a similar survey (see the table, right side) of 132 visitors to the Great Ape Trust of Iowa (GATI) (Des Moines, IA) in 2006 (2). There, only 72% of respondents thought chimpanzees were endangered (compared with 94% for gorillas and 92% for orangutans) (for CvG, $\chi = 22.53$, $P < 0.001$; for CvO, $\chi = 17.21$, $P < 0.001$). Of those who did not believe chimpanzees were endangered, 30% justified their response by noting how often they see chimpanzees in the media and as pets. Given the reality facing chimpanzees in the wild—current estimates are that populations could go extinct in the next several decades (3)—such data highlight the importance of accurately representing chimpanzees

Depictions of chimpanzees as caricatures can lead people to think these animals are not endangered and is a problem for conservation and welfare efforts.



SURVEY RESPONSES

	RCAA		GATI	
	Yes	No	Yes	No
Chimpanzees	665	333	95	37
Gorillas	953	45	124	8
Orangutans	911	87	121	11

and other apes in popular media, particularly by professional publications such as *Science*.

The inappropriate portrayal of great apes in advertisements undermines the scientific, welfare, and conservation goals that we and many readers work hard to achieve (4). Respected organizations such as AAAS must take a leadership role in promoting ethically sound practices not only in research they promote, but in fields as diverse as public relations and marketing. Together with like-minded organizations such as the International Society of Primatologists (IPS) and the Association of Zoos and Aquariums (AZA), we can make progress in shifting the perception of chimpanzees as frivolous subhumans that are not in danger of extinction to more scientifically accurate characterizations of our closest relatives that stir interest, respect, and conservation efforts.

References and Notes

1. Although some states have strict permitting regulations, grandfather clauses, and other allowances, we have identified only six states in which it is impossible for a private citizen (nonzoo, nonresearch) to own a chimpanzee. Although there is progress on this front in several states, at least 21 states have no or minimum permitting requirements at all (S.R.R. and S. Baekler, personal communication).
2. Materials and methods are available as supporting materials on Science Online.
3. C. Tutin et al., *Regional Action Plan for the Conservation of Chimpanzees and Gorillas in Western Equatorial Africa* (Center for Applied Biodiversity Science, Conservation International, Washington, DC, 2005).
4. In addition to these potentially broad effects on conservation attitudes, the use of chimpanzees by the entertainment industry deserves attention if only for individual welfare considerations. Although these animals are technically regulated by the Animal Welfare Act, enforcement of these regulations is difficult, and many chimpanzees are subject to suboptimal rearing, housing, and care, with short- and long-term welfare consequences.

Supporting Online Material

www.sciencemag.org/cgi/content/full/319/5869/1487/DC1

¹Lester E. Fisher Center for the Study and Conservation of Apes, Lincoln Park Zoo, Chicago, IL 60614, USA. ²Chimpanzee Species Survival Plan (SSP) of the Association of Zoos and Aquariums (AZA), Silver Spring, MD 20910, USA. ³Cleveland Metroparks Zoo, Cleveland, OH, and Case Western Reserve University, Cleveland, OH 44109, USA. ⁴Committee on Evolutionary Biology, University of Chicago, Chicago, IL 60637, USA. ⁵Zoo Atlanta, Atlanta, GA 30315, USA. ⁶Ape Taxon Advisory Group (TAG) at AZA, Silver Spring, MD 20910, USA. ⁷Duke University, Durham, NC 27708, USA. ⁸Great Ape Trust of Iowa, Des Moines, IA 50320, USA. ⁹Krasnow Institute for Advanced Study, Fairfax, VA 22030, USA. ¹⁰The Jane Goodall Institute, Arlington, VA 22203, USA.

*Author for correspondence. E-mail: ross@lpzoo.org

PLANETARY SCIENCE

Observing Our Origins

Fred J. Ciesla

Planetary systems are born around young stars and grow from vast clouds of dust and gas called protoplanetary disks. Models predict that as our own solar system's protoplanetary disk evolved, the dust and gas pushed each other around while constantly being stirred and jolted by magnetic fields and gravitational torques. The resulting mixing and motion set the chemical compositions of the planetesimals that formed and from which planets eventually grew. Although evidence for mixing is found in objects in our solar system, such as primitive meteorites, questions

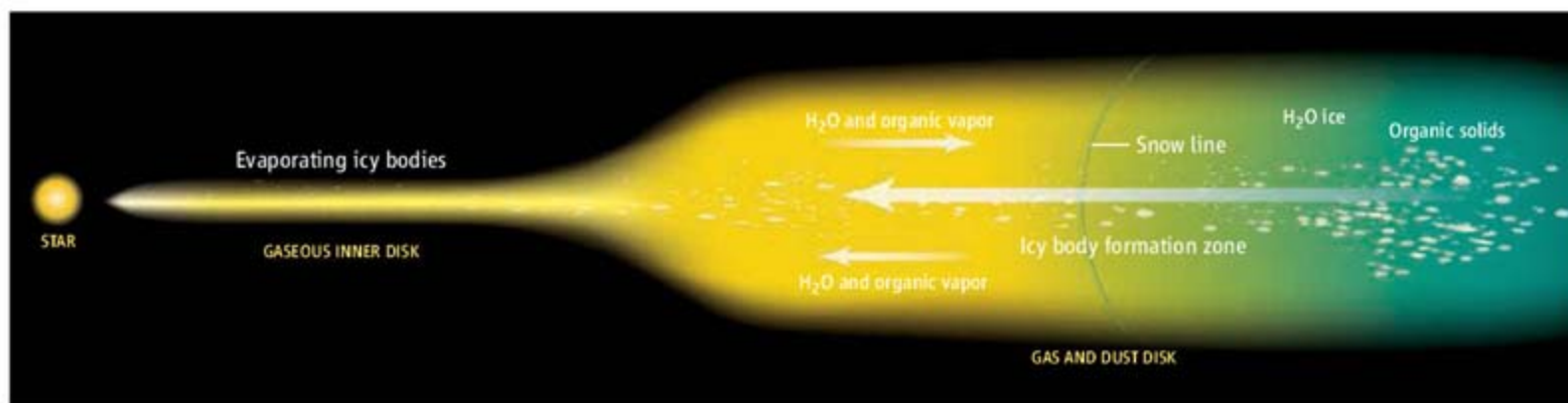
regions of the disk, they vaporize, with the vapor then either continuing inward with the flow of disk material, albeit at a slower rate, or diffusing outward where it would freeze out at lower temperatures and be incorporated into other solids only to repeat the journey again (see the figure). The whole process acts as a chemical conveyor belt, delivering materials from one part of the disk to another, where they can accrete or undergo chemical reactions.

Models of this transport and cycling have largely focused on water, owing to its high

Infrared observations of water and organic materials will help astronomers determine how our solar system formed.

explanation for the range of isotopic ratios seen in chondritic materials (6, 7).

Although agreements of this type between the models and the chondrites hint that we are beginning to understand how our solar system formed, they are far from definitive. Alternate models have been proposed and have equal success in explaining the properties of chondritic materials. For example, Clayton (8) suggests that the differences in oxygen properties arise because chondritic materials formed at high temperatures near the Sun and were then tossed outward by strong jets. Identifying



Birthplace of the planets. The temperature of a protoplanetary disk decreases with distance from the star, so some compounds “freeze out” of the gas beyond a certain point in the disk. For example, water condenses to form water ice beyond the “snow line.” Models predict that water ice boulders drifting into the warm inner region of the disk, where terrestrial planets are expected to form, supply this region

with water vapor that would otherwise be lost. Organic compounds could be delivered in a similar manner. Carr *et al.* report signatures of these materials in the inner disk of the star AA Tau, providing observational support for such models. These processes may be responsible for the chemical and isotopic variations observed in the chondritic meteorites that formed within our own protoplanetary disk.

remain about the details of the processes responsible and whether this mixing was common in other protoplanetary disks. On page 1504 of this issue, Carr *et al.* (1) report observations of the disk around the star AA Tau that suggest that we will soon be able to address these questions.

The gas and solids in a protoplanetary disk drift inward with time, representing the final stages of growth for the star (2). On their journey inward, the solids collide with one another and stick together. From the time they grow from the submicrometer-sized grains originally inherited from the parent molecular cloud to the kilometer-sized bodies that eventually accrete to form planets, solids migrate inward slightly faster than the gas (3). As bodies drift into hotter

abundance in the solar nebula (4). Initially, the influx of water ice to the warm inner nebula is predicted to have been large, leading to an increase in the relative abundance of water vapor in the gas phase. With time, the influx decreased as water was locked up in immobile planets and comets, allowing the vapor to be redistributed without being replenished, leading to a decrease in the abundance of water in the inner disk.

Because water is a major oxidizing agent, variations in its abundance would be recorded in the mineralogy of the rocks that formed in the solar nebula. Indeed, chondritic meteorites—relatively unaltered remnants from the solar nebula—contain minerals that formed in environments ranging from highly oxidizing (increased abundance of oxygen) to very reducing (depleted oxygen) (5). In addition, photochemical effects could enrich water with the heavy isotopes of oxygen, providing an

which of these models is correct is critical to furthering our understanding of how planetary systems form because it has implications on other issues, ranging from the origin of cometary grains to the manner by which giant planets form. Unfortunately, researchers have been unable to settle the debate by examining the chondrites alone.

This is what makes the results of Carr *et al.* so exciting. With data from the Spitzer Space Telescope, they identified spectral emission features of water and basic organic molecules from the gas in the disk around AA Tau within 3 AU (one AU or astronomical unit is the distance from Earth to the Sun) of the central star, approximately where the chondrites formed in our solar nebula. Carr *et al.* infer abundances greater than predicted in stagnant disks, suggesting a dynamically active disk where inward drifting, volatile-rich boulders replenish water and organic molecules. Water emis-

Department of Terrestrial Magnetism, Carnegie Institution of Washington, Washington, DC 20015, USA. E-mail: fciesla@ciw.edu

sion features have also been reported for the disks around the stars MWC480 (9) and SVS 13 (10), although these studies only probed distances <0.3 AU from the central star. Interestingly, water appears to be depleted in SVS 13 relative to what is predicted in stagnant disk models (10). The variation of observed water abundances in these disks mirrors that which has been inferred for our own solar nebula.

To date, these observations do not distinguish which of the models developed for our solar nebula is correct but rather lend support to recent models for the dynamic evolution of water and other volatiles in protoplanetary disks. However, as the techniques used by Carr *et al.* are applied to other disks, correlations between their chemical compositions and their physical properties can be

identified. Models for water evolution predict that the enhancement of water in inner disks should be followed by periods of depletions, so systematic variations with age are expected. Also, larger disks would provide more water ice to drift inward and thus would produce greater enhancements in the inner disk. Searching for such correlations will thus allow us to test models developed for our own solar nebula and determine whether it evolved in a similar way as other disks in our galaxy or if, instead, our planetary system is the result of one or multiple unique circumstances. Right now, these new results, combined with the discovery of high temperature grains in comets (11) and in the outer regions of protoplanetary disks (12), suggest that the manner by which our solar system formed may have been the rule.

References

1. J. S. Carr *et al.*, *Science* **319**, 1504 (2008).
2. C. P. Dullemond, D. Hollenbach, I. Kamp, P. D'Alessio, in *Protostars and Planets V*, B. Reipurth, D. Jewitt, K. Keil, Eds. (Univ. Arizona Press, Tucson, 2007), pp. 555–572.
3. S. J. Weidenschilling, *Mon. Not. R. Astron. Soc.* **180**, 57 (1977).
4. F. J. Ciesla, J. N. Cuzzi, *Icarus* **181**, 178 (2006).
5. A. N. Krot, B. Fegley Jr., K. Lodders, H. Palme, in *Protostars and Planets IV*, V. Mannings, A. P. Boss, S. S. Russell, Eds. (Univ. Arizona Press, Tucson, 2000), pp. 1019–1054.
6. H. Yurimoto, K. Kuramoto, *Science* **305**, 1763 (2004).
7. J. R. Lyons, E. D. Young, *Nature* **435**, 317 (2005).
8. R. N. Clayton, *Nature* **415**, 860 (2002).
9. J. A. Eisner, *Nature* **447**, 562 (2007).
10. J. S. Carr, A. T. Tokunaga, J. Najita, *Astrophys. J.* **603**, 213 (2004).
11. D. Brownlee *et al.*, *Science* **314**, 1711 (2006).
12. R. van Boekel *et al.*, *Nature* **432**, 478 (2004).
13. R. van Boekel, *Nature* **447**, 535 (2007).

10.1126/science.1155858

SYSTEMS BIOLOGY

Customized Signaling Circuits

Peter M. Pryciak

For nearly three decades, cell biologists have labored to identify and dissect the elaborate intracellular signaling pathways that control cellular responses to external stimuli. The emerging field of “synthetic biology” now seeks to move beyond mere understanding of these existing biological systems, and to begin exploiting the acquired knowledge for new purposes such as creating custom-configured signal transduction pathways (1–3). Much as an engineer assembles new electronic circuits from a toolbox of pre-

existing parts, the study by Bashor *et al.* on page 1539 in this issue (4) modifies and re-connects components of a well-characterized cellular signaling pathway to reshape fundamental input-output processing behaviors such as temporal dynamics and dose response.

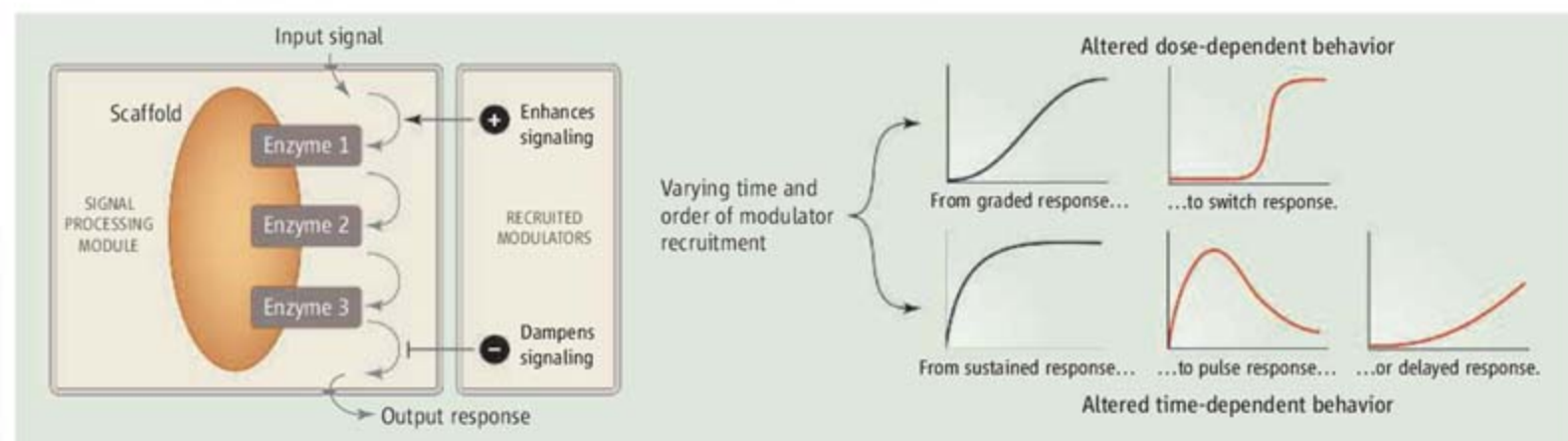
The system chosen for modification is the signaling pathway that responds to mating pheromones in the budding yeast *Saccharomyces cerevisiae*. Because this pathway has long been a model for eukaryotic signal transduction (5), the depth of knowledge and the ease of experimental manipulations make it an ideal system for testing new theories of pathway engineering.

In principle, two general strategies can be

Altering cellular behaviors can be achieved through a synthetic approach by refashioning signaling circuitry.

used to alter signaling circuitry: a bottom-up approach involving de novo design of proteins with new properties (e.g., new interactions, substrate specificities, or kinetic parameters), or a modular approach in which existing proteins are co-opted as parts to be connected in new ways. Bashor *et al.* follow the latter scheme, which exploits the modular property of many natural signaling proteins (6). At the core of this effort lies a “scaffold” protein called Ste5, which serves as an assembly platform for a series of sequentially acting enzymes (protein kinases) that propagate signals through the pathway (7). The role of scaffold proteins as central signal processing hubs makes them a natural choice as the framework upon which to

Department of Molecular Genetics and Microbiology, University of Massachusetts Medical School, Worcester, MA 01605, USA. E-mail: peter.pryciak@umassmed.edu



Changing behavior. The response of a cellular signaling pathway to a stimulus can be altered with positive and negative modulators. When such modulators are recruited to the scaffold protein in specific temporal sequences, through the use

of feedback loops that control their expression and competitor recruitment sites that act as binding sinks, the time or dose dependence of the signaling response can be adjusted to adopt a variety of useful circuit behaviors.

append additional regulatory input. Indeed, previous work indicated that modified scaffolds can alter the flow of signaling between alternate pathways (8, 9). Bashor *et al.* build on these past efforts in a comprehensive and systematic way, generating multiple new layers of control over signaling dynamics.

The authors use heterodimerizing protein interaction motifs called leucine zippers to recruit to the scaffold protein additional positive or negative modulators of pathway signaling. Alone, recruitment of these modulators simply enhances or dampens signaling. To generate more sophisticated behaviors, however, the timing of their recruitment was varied in two ways: by expressing modulators from promoters that themselves are regulated by the signaling pathway, thus generating feedback loops, and by forcing the modulators to compete for access to the scaffold with nonfunctional "decoy" molecules, thus generating delayed action. Different permutations of these variables yielded different effects on either temporal or dose-response behaviors. For example, if expression of the negative modulator is induced by the pathway, but must first saturate a constant number of high-affinity decoy binding sites before it can bind to the scaffold, the pathway is converted from one that shows sustained activation to one that shows a "pulse" of activation followed by a sharp decline (see the figure). A reciprocal arrangement, in which a preexisting negative modulator must be displaced by a decoy protein whose expression is induced by the signaling pathway itself, causes the response to be delayed, rather than immediate. Yet another configuration alters the dose dependence of the pathway, converting it from a graded, rheostat-like response to a sharply sensitive, switchlike response. A related recent study using the same system showed that by expressing pathway components from a promoter that is itself regulated by the same signaling pathway, a positive-feedback loop can be established that maintains signaling even after the stimulus is removed, thus converting the pathway from reversible to irreversible (10).

The results of such tinkering illustrate several points. They test whether our concepts about signaling mechanisms are correct. Indeed, the observed results clearly emphasize how colocalization of signaling proteins can play a critical role in shaping pathway behavior. In addition, the observed variations in signaling behavior mimic those in nature (e.g., transient versus sustained, or graded versus switchlike) and, hence, suggest how they could have arisen by evolutionary swapping of promoters or protein-binding sites. The results also show that signaling dynamics can be successfully reengineered by using rational approaches.

One goal of synthetic biology is to establish a set of standard biological parts that can be connected in multiple combinations to accomplish various objectives (3). Although some tools developed by Bashor *et al.* may seem pathway specific, it is conceivable that the entire signaling cascade, along with the modifications that confer specific circuit behaviors, might serve as a transferable module that can be connected to different inputs and outputs. Indeed, because a primary output of this pathway is the regulation of gene expression, essentially any gene can be placed under pathway control by simply providing it with the proper promoter. Furthermore, the activating stimulus can be altered either by additional changes to the scaffold (8, 9) or by replacing the upstream receptor of the stimulus (11). Therefore, future engineers might potentially generate an extraordinary variety of signaling circuits by mixing and matching one choice from each category: an input, an

output, and a signal-processing module. The broad lesson of the modular approach is that complex behaviors do not necessarily require highly evolved proteins, but can be developed from the gradual layering of regulators and connections. These early studies are just the tip of the iceberg in what is likely to become a rapidly accelerating field.

References

1. T. Pawson, R. Linding, *FEBS Lett.* **579**, 1808 (2005).
2. D. Sprinzak, M. B. Elowitz, *Nature* **438**, 443 (2005).
3. D. A. Drubin, J. C. Way, P. A. Silver, *Genes Dev.* **21**, 242 (2007).
4. C. J. Bashor, N. C. Helman, S. Yan, W. A. Lim, *Science* **319**, 1539 (2008).
5. L. Bardwell, *Peptides* **26**, 339 (2005).
6. R. P. Bhattacharyya *et al.*, *Annu. Rev. Biochem.* **75**, 655 (2006).
7. N. Dard, M. Peter, *Bioessays* **28**, 146 (2006).
8. K. Harris *et al.*, *Curr. Biol.* **11**, 1815 (2001).
9. S.-Y. Park, A. Zarrinpar, W. A. Lim, *Science* **299**, 1061 (2003).
10. N. Ingolia, A. W. Murray, *Curr. Biol.* **17**, 668 (2007).
11. S. J. Dowell, A. J. Brown, *Recept. Channels* **8**, 343 (2002).

10.1126/science.1156414

MATERIALS SCIENCE

The New Diamond Age?

Paul W. May

After the hype, what realistic applications might synthetic diamond films have in the near future?

Diamonds were prized for their scarcity for centuries, and they remain a symbol of wealth and prestige to this day. Apart from their appeal as gemstones, diamonds have remarkable physical properties. Diamond is the hardest known material, has the highest thermal conductivity at room temperature, is transparent over a wide range of wavelengths, is the stiffest and least compressible material, and is inert to most chemical reagents. It is thus not surprising that diamond has been referred to as the ultimate engineering material. Here I highlight some of the exciting new areas where the use of artificial diamond in the form of thin films or coatings may find realistic wide-scale applications in the next few years.

Artificial diamond was first fabricated in the laboratory in the 1950s by the high-pressure, high-temperature growth technique. This method has been used to produce small synthetic diamond crystals, which are used for industrial processes such as cutting and machining mechanical components and for polishing and grinding of optics.

School of Chemistry, University of Bristol, Bristol BS8 1TS, UK. E-mail: paul.may@bris.ac.uk

In the late 1980s, a new method of making diamond was developed (1). In the chemical vapor deposition (CVD) method, a gas-phase chemical reaction above a solid surface results in deposition onto that surface. For diamond, the process gas is usually a mixture of 99% H₂ and 1% CH₄, activated by a hot (2000°C) metal filament or a microwave plasma. A substrate temperature above 700°C ensures formation of diamond rather than amorphous carbon. Apart from diamond itself, the most common substrate material is silicon; researchers now regularly grow polycrystalline diamond films to thicknesses from micrometers to millimeters on standard Si wafers. Adding a boron-containing gas to the process mixture allows the diamond film to become boron-doped, giving it controllable p-type semiconducting properties.

In the early 1990s, the rapid progress in this field led to speculation that diamond would become the next-generation ideal semiconductor and spark a new "diamond age" for electronics and mechanical components. This technological promise has not yet been realized. Was it all just hype? And what are the realistic applications for CVD diamond in the short to middle term?

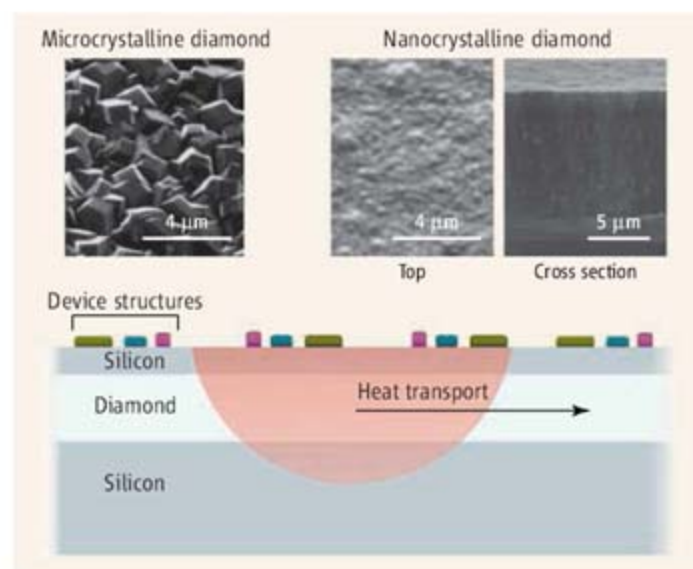
Some of the major problems associated with diamond CVD have now been solved, but many hurdles remain. First, it is difficult to make large-area, single-crystal diamond substrates. Most CVD diamond films are polycrystalline, and the presence of graphitic material at the grain boundaries between the diamond crystallites degrades the material properties. Second, diamond films grow at a rate of only a few micrometers per hour; unless methods are found to speed up growth by a factor of 10 or more, diamond films will remain prohibitively expensive.

Third, the high deposition temperature severely limits the substrates on which diamond can be deposited. Ultra-nanocrystalline diamond with grain diameters below 10 nm can be grown at 400°C (2), but most CVD diamond requires temperatures above 700°C. Thus, low-melting-point materials such as glass, plastics, and aluminum are difficult, if not impossible, to coat with diamond by standard CVD methods. Ferrous materials are also problematic, because carbon is highly soluble in iron at these temperatures.

Fourth, doping diamond to make a viable n-type semiconducting material has proved elusive. Phosphorus has been used as a dopant, but the electronic characteristics are not good enough for most devices (3). Without an n-type material, diamond electronics will be limited to simple p-type devices such as sensors and detectors. Even if a suitable n-type dopant were discovered, it is doubtful that the resulting material would have high enough carrier mobilities to rival high-performance room-temperature electronics made with silicon. However, for high-power, high-temperature electronics, the properties of diamond reign supreme. High-value specialized applications such as sensors in car or jet engines where temperatures can be above 300°C, or fast high-power switches for train doors or heavy machinery, are areas where diamond may have a substantial impact on the market.

Despite the problems mentioned above, thin, polycrystalline diamond films currently find small-scale specialized applications. They are, for example, used as hard and wear-resistant coatings, cold cathode sources, heat spreaders, microwave and infrared windows, micro- and nanoelectromechanical systems, electrolytic water purifiers, ultraviolet sensors, electrodes for electrochemical processes, and loudspeaker tweeters (4).

A promising new application stems from the recent discovery that the usually inert surface of diamond can be functionalized by chemically replacing the hydrogen atoms that terminate the carbon bonds with other, more



The next generation of integrated circuits? In Si-diamond-Si sandwich structures, the heat generated by the devices (red semicircle) diffuses into the diamond layer, which rapidly transports the heat to a heat sink at the chip edge, preventing overheating. The diamond layer may be microcrystalline diamond (top left), which must be polished before use, or nanocrystalline diamond [shown in top view (top middle) and cross section (top right), with film thickness 10 μm], which conducts heat less well but does not require polishing.

reactive ligands. Attaching a hydrocarbon chain terminating in an amine group allows selected proteins or DNA fragments to be tethered to the diamond surface (5). When immersed in a liquid solution containing biomolecules, the tethered protein/DNA will selectively grab hold of its natural “partner” (for example, a coenzyme or complementary DNA strand). This uptake can be monitored electrochemically or via a fluorescent marker. This approach could be used to make cheap, reusable biosensors for use in genetic screening, disease diagnosis, or day-to-day monitoring of diabetes and other conditions.

The extremely high thermal conductivity of diamond may result in one of the first large-scale applications for CVD diamond films. With the increasing miniaturization of silicon-integrated circuits comes the major problem of heat dissipation. To overcome this, chip manufacturers and diamond companies such as sp³ Diamond Technologies Inc. (6) are beginning to consider replacing the standard Si wafer substrate with a composite wafer comprising a Si-diamond-Si sandwich. To make this wafer, a CVD diamond layer 10 to 20 μm thick is grown onto a standard Si wafer. Next, a thin (<5 μm) layer of Si is deposited. The device components are deposited onto the top Si layer and patterned using standard photolithography and etching processes (see the figure).

Although this sounds relatively straightforward, some serious technological problems

remain. These include devising methods to deposit diamond films uniformly onto 300-mm-diameter Si wafers, and then polishing the surface to nanometer smoothness, while avoiding bowing of the wafer that might arise from thermal-expansion mismatch between the diamond and Si. If these challenges are overcome, it is likely that within 5 years, most PC microprocessors may be fabricated on such a diamond-sandwich wafer.

The expense and difficulties of making such complex composite wafers would not be necessary if large single-crystal diamond wafers were commercially available at affordable prices. Progress in growing single-crystal diamond has been slow, but breakthroughs in understanding

the growth chemistry in the past 5 years are beginning to yield results. The UK-based company Element Six now makes and supplies single-crystal diamond substrates, although the largest are only about 5 μm by 5 μm (7). Ho *et al.* have grown single-crystal CVD diamond gemstones 6.7 mm in diameter and 12 mm high (8). The crystals were intended primarily for use in diamond anvil cells, but the same technology can be used to make commercial gems. Companies such as Apollo Diamond (9) are beginning to exploit these advances for growing commercial gemstones.

In hindsight, the enthusiastic predictions of the 1990s for diamond were perhaps too ambitious, but the story is not over yet. As well as being essential for a growing number of small but lucrative niche applications, CVD diamond may yet play a crucial role in more mainstream emerging technologies.

References

1. P. W. May, *Philos. Trans. R. Soc. London Ser. A* **358**, 473 (2000).
2. D. M. Gruen, O. A. Shenderova, A. Ya. Vul', Eds., *Synthesis, Properties and Applications of Ultrananocrystalline Diamond* (Springer, New York, 2005).
3. R. Kalish, *Diamond Relat. Mater.* **10**, 1749 (2001).
4. K. E. Spear, J. P. Dismukes, *Synthetic Diamond: Emerging CVD Science and Technology* (Electrochemical Society, Pennington, NJ, 1994).
5. C. E. Nebel *et al.*, *J. R. Soc. Interface* **4**, 439 (2007).
6. sp³ Diamond Technologies Inc. (www.sp3inc.com).
7. Element Six (www.e6cvd.com).
8. S. S. Ho, C. S. Yan, Z. Liu, H. K. Mao, R. J. Hemley, *Ind. Diamond. Rev.*, **28** (January 2006).
9. Apollo Diamond Inc. (www.apollosdiamond.com).

PHYSICS

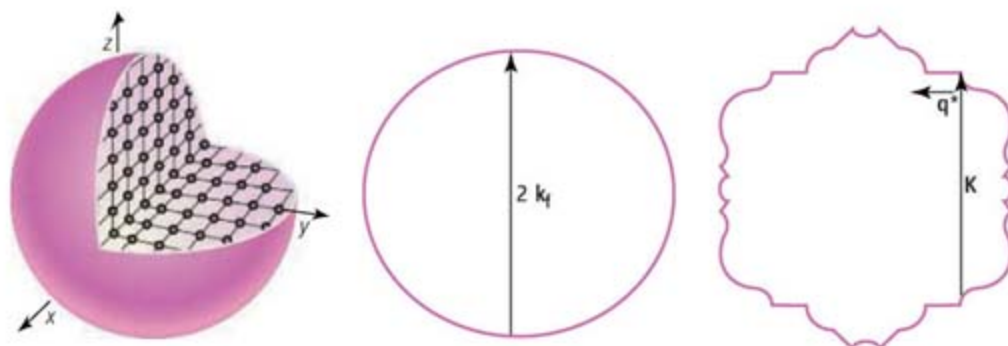
This Coincidence Cannot Be Accidental

Douglas J. Scalapino

For the past two decades, physicists have been obsessed with figuring out how cuprates can be superconducting at high temperatures. Yet on page 1509 of this issue, Aynajian *et al.* (1) report neutron scattering measurements indicating that long-known superconductors such as lead and niobium still have plenty of puzzling features. The data imply that either there is new physics not contained in the accepted theory of superconductivity (2, 3) or there is some aspect of this theory that we have not yet recognized. Either way, it is remarkable to learn something new about elemental metallic superconductors, which we thought were understood more than 40 years ago.

Lattice vibrations called phonons are the key here; phonons in metals hold electrons together in pairs, and these pairs then form a current that flows without resistance. Aynajian *et al.* show that as the frequency of the phonons increases, their lifetime exhibits dips at several frequencies ω_{KA} called “Kohn anomalies” (4). What is remarkable is that the energy of the phonon $\hbar\omega_{KA}$ (where \hbar is Planck’s constant divided by 2π) associated with one of these Kohn anomalies appears to coincide with the measured binding energy $2\Delta(0)$ of superconducting electron pairs [where $\Delta(0)$ is the zero-temperature superconducting energy gap]. The energy gap and the energies of the Kohn anomalies are of course different for lead and niobium, but the equality of the phonon energy associated with one of the Kohn anomalies and the pair binding energy $2\Delta(0)$ is observed for both materials. As the authors remark, “this coincidence cannot be accidental.”

To appreciate this surprising equality, we need to consider the differences in the physics responsible for a Kohn anomaly and for superconducting pairing. The Kohn anomalies arise from and reflect the way in which the electrons fill the energy levels in a metal. For example, if we ignore the ions, the one-electron states of the conduction electrons are plane waves. Representing the metal as a periodic cube of side L , the allowed values of the momentum are quantized such that $\mathbf{p} = \hbar\mathbf{k}$



Resolving an anomaly. (Left) Cutaway view of a spherical Fermi surface that separates the filled \mathbf{k} -states (shown as dots) from the empty ones. (Middle) A cross section of the Fermi sphere that encloses the occupied one-electron \mathbf{k} -states of a noninteracting electron gas. The solid line denotes the Fermi surface and $2\mathbf{k}_f$ is the wave vector of a Kohn anomaly. (Right) Schematic diagram of the (110) cross section of the Pb Fermi surface. Here \mathbf{K} denotes one of the reciprocal lattice vectors, and a phonon with wave vector \mathbf{q}^* has a Kohn anomaly. The electron-phonon coupling to transverse acoustic phonons is initiated when \mathbf{q} exceeds \mathbf{q}^* . Here the electron scattering process involves a reciprocal lattice wave vector \mathbf{K} and is called an *umklapp* scattering.

with allowed wave vectors $\mathbf{k} = (2\pi/L)(n_x, n_y, n_z)$. Here the n_i are integers, and one can specify the one-electron states by points in a three-dimensional k -space.

Electrons obey the Pauli exclusion principle, which prevents two of them from occupying the same state. The ground state is characterized by having all the allowed k -space points inside a sphere of a given radius k_f filled and the states outside empty (see the left panel of the figure). The momentum up to the filled radius is called the Fermi momentum, and the surface of the k -space sphere that separates the occupied from the unoccupied states is called the Fermi surface (a slice through the surface is shown in the middle panel). When the potential from the ionic lattice is taken into account, the one-electron states are altered and the Fermi surface changes, as illustrated for Pb (see the right panel of the figure). The reciprocal lattice vectors \mathbf{K}_i are equal to $2\pi/a_i$, where a_i describes the unit cell of the Pb lattice. The reflection of the electron waves from the static ionic lattice gives rise to the geometry of the Fermi surface characteristic of a given material.

The conduction electrons of a metal surround and screen the electric field of the ionic lattice. The restoring force of this field enters in determining the frequency and lifetimes of the phonons. As shown by Kohn (4), the ability of the electrons to screen the ionic electric field decreases suddenly at certain phonon wavelengths λ , determined by the geometry of the Fermi surface. This leads to a change in the

Metal superconductors such as lead and niobium were thought fully understood 40 years ago, but they have now presented some fresh puzzles.

phonon frequency and an increase in its dissipation, which is reflected as a dip in its lifetime. This is the Kohn anomaly. Wave vectors $q_{KA} = 2\pi/\lambda$ associated with these wavelengths connect parts of the Fermi surface that have parallel tangents. Two examples of such wave vectors, $q_{KA} = 2k_f$ and $q_{KA} = q^*$, are shown in the figure. The phonon frequency ω_{KA} at which the Kohn anomaly occurs is determined by q_{KA} . For the low-frequency anomaly discussed by Aynajian *et al.*, $\omega_{KA} = v_s^t q_{KA}$, where v_s^t is the transverse acoustic sound velocity, so that $\hbar\omega_{KA} = \hbar v_s^t q_{KA}$.

In contrast to the phonon Kohn anomaly energy $\hbar\omega_{KA}$, which is determined by the geometry of the one-electron Fermi surface and the velocity of sound in the normal nonsuperconducting state, the binding energy $2\Delta(0)$ involves pairs of electrons and is a property of the superconducting state. It is equal to the energy required to break apart a superconducting pair. Although the effective pairing interaction involves phonons, it is difficult to see any direct connection between $\hbar\omega_{KA}$ and $2\Delta(0)$.

As an analogy, imagine watching rainfall and snow totals for the Lake Tahoe area in the western United States in order to estimate how high the lake will be in the summer. We notice that when the official rainfall season ends in May, Lyon County, a neighboring agricultural area in Nevada, reports 38.6 inches of rain. Then late in June, after the spring thaw, we get an announcement that Lake Tahoe is once again running over the spillway of the small dam at

Department of Physics, University of California, Santa Barbara, CA 93106, USA. E-mail: djs@physics.ucsb.edu

the northeast end of the lake and had risen 38.6 inches since the time it last spilled on 1 November. Checking records from previous years, we observe that when the lake spills in June, its reported rise is always equal to the annual rainfall of Lyon County. Although some winter storms that bring precipitation to the Tahoe area move on to Lyon County, there clearly must be something else at work. Indeed, in this fictitious story, we discover that there is a local law that during the summer months the dam must be lowered to keep the Truckee River flowing from Tahoe to Lyon County. On 1 November, the spillway is to be raised because the winter rains will keep the river flowing. Then on 1 June, the spillway is to be lowered to a height above its 1 November level equal to the annual rainfall recorded in Lyon County. Thus, it is not a coincidence but a law that locks the rise of Lake Tahoe to the rainfall in Lyon County.

Similarly, although the polarization created by the conduction electrons and the response of the ionic lattice contribute to determining both the Kohn anomaly and the pair binding energy, there must be something else at work that locks $2\Delta(0)$ to $\hbar\omega_{KA}$. As noted, this could mean that there is some new physics that is not captured by the Eliashberg formulation (3). Aynajian *et al.* raise the possibility that charge density or spin density correlations may limit the growth of the superconducting energy gap and lead to the observed convergence of $2\Delta(0)$ and $\hbar\omega_{KA}$.

Alternatively, I believe that the Eliashberg theory contains an explanation for the experimental findings, provided that one takes into account the change of the pairing interaction that occurs in the superconducting state. Suppose the Kohn anomaly is associated with the scattering process illustrated in the right panel of the figure. Then, at energies greater than $\hbar\omega_{qs}$, there will be a sudden increase in the coupling of electrons to the transverse acoustic mode because K is parallel to the polarization (5, 6). Then in the superconducting state, if $2\Delta(0)$ is close in energy to $\hbar\omega_{qs}$, it may lock on to it. That is, for $2\Delta(0) < \hbar\omega_{qs}$, the change in the electron polarization caused by the superconducting gap will push phonon spectral weight above $\hbar\omega_{qs}$, where it can couple via an *umklapp* scattering process (from the German word for “flip over”) shown in the right panel. This will lead to an increase in $2\Delta(0)$. Alternatively, if $2\Delta(0)$ increases above $\hbar\omega_{qs}$, the electron polarization will push phonon spectral weight below $\hbar\omega_{qs}$, where it will not couple, causing a decrease in $2\Delta(0)$.

Although this locking can only occur if $\hbar\omega_{KA}$ is reasonably close to the value $2\Delta(0)$ would have had in the absence of Kohn anom-

alies, it is not a complete accident. Furthermore, it can be understood within the traditional theory of superconductivity. It appears from the electron tunneling data that an *umklapp* Kohn anomaly in the transverse phonon is indeed present in Pb (5, 6). One needs to see whether a similar feature can be observed in the Nb tunneling spectrum. A further test would be a careful measurement of the size of the superconducting gap on different parts of the Fermi surface. One only needs the “locking” on the Fermi surface regions associated with the *umklapp* Kohn anomaly.

ARCHAEOLOGY

Standing on the Shoulders of Giants

Fiona Coward

Archaeological studies are helping to understand how humans acquired the ability for cultural transmission.

How and why did humans learn to learn? A variety of disciplines have recently provided important insights into the basic mechanisms that underpin cultural transmission. Archaeology is now beginning to place these insights in a chronological framework that will help to understand when and why these mechanisms evolved among our ancestors.

The extent to which human behaviors and knowledge are culturally transmitted within and between generations has long been considered a defining feature of our species. Parts of the behavioral repertoires of many other animals—from ants to dolphins—are neither determined by genetics nor individually acquired but learned from members of the same species. For example, the manufacture and/or use of material objects among some of our closest primate relatives are group-specific and persist between generations (1, 2). However, the diversity and complexity of learned behaviors among humans by far outstrip anything known in other species; in addition, human culture is cumulative in a way that other species' socially learned repertoires are not (3, 4).

A basic capacity for motor imitation—the mirror neuron system, which automatically maps the observed actions of others onto one's own motor system—is part of our primate heritage (5). But, although

References and Notes

1. P. Aynajian *et al.*, *Science* **319**, 1509 (2008); published online 21 February 2008 (10.1126/science.1154115).
2. J. Bardeen, L. N. Cooper, J. R. Schrieffer, *Phys. Rev.* **108**, 1175 (1957).
3. G. M. Eliashberg, *Sov. Phys. JETP* **16**, 780 (1963).
4. W. Kohn, *Phys. Rev. Lett.* **2**, 393 (1957).
5. D. J. Scalapino, in *Superconductivity*, R. Parks, Ed. (Dekker, New York, 1969), vol. 1, chap. 10, fig. 30.
6. W. L. McMillan, J. M. Rowell, *Phys. Rev. Lett.* **14**, 108 (1965).
7. I thank T. Devereaux, W. Harrison, S. Kivelson, and R. Martin for helpful discussions, S. Johnson for help with the figure, and the Stanford Institute for Theoretical Physics for its support.

10.1126/science.1155859

other species may learn behaviors and even act in such a way as to facilitate their offspring's learning, only humans are known to teach, which involves actively correcting learners (4, 6). Furthermore, although primates are capable of complex interactions



Bead making at Makuri Village, Namibia.

Department of Geography, Royal Holloway University of London, Egham TW20 0EX, UK. E-mail: Fiona.Coward@rhul.ac.uk



Learning traditional Micronesian navigation skills.

with one another or with objects, they do not seem to engage in relations with objects and with other individuals at the same time (6). The suggestion is that the human mirror neuron system may allow us to go beyond imitating the observed motor acts of others to infer their intentions and perhaps even their states of mind (5)—perhaps the prerequisite for true imitation and cumulative cultural transmission.

However, neither humans nor primates are born fully fledged imitators or mind readers; the necessary cognitive and motor systems take time to mature. The fossil record documents an extension of the hominin developmental period relative to that of primates: Estimates of brain size at birth, coupled with analysis of the relative development of teeth and bones of juveniles, demonstrate the birth of less-developed infants and longer, slower growth of both brains and bodies (7). Stone tools appear in the archaeological record from at least 2.5 million years ago (8), roughly at the same time as the earliest known specimens of *Homo*, documenting sufficient social and technical skills for the habitual targeting of higher-quality foods requiring more processing to extract, such as meat. This dietary shift in turn made it easier to provide for the longer developmental period, which required the involvement of more than one adult—an indication of more complex and longer-lasting social relationships.

This expanded period of development and maturation of the brain thus occurred in increasingly rich social and cultural environments, with longer-lived social relations

facilitating the transmission of more and more complex cultural skills (4)—many craft skills practiced by modern humans take several years of intensive teaching to master, often in childhood (see the first photo). Neuroimaging studies of the acquisition of toolmaking skills (9) and modeling of early hominin social systems based on those of extant primates (10) are fleshing out our understanding of the basic cognitive mechanisms for motor imitation, learning, and sociality. Their social and cultural flexibility allowed hominins to colonize new and unfamiliar ecosystems and to develop the bewildering diversity of cultural traits visible today and in the archaeological record as they spread across the globe.

Archaeology can add to the debate by investigating specific patterns of cultural transmission among and between prehistoric populations. The patterns formed by the geographic and temporal distribution of material culture in the archaeological record results from the dissemination of the relevant behaviors between individuals. Thus, the process can be modeled (much as epidemiologists model the transmission of disease) to investigate the social and cultural factors that influence how learned skills spread into population-wide distributions (3). Application of such models to the archaeological record has provided insights into such puzzles as the loss of various technologies in Holocene Tasmania, including the manufacture of composite tools and of cold-weather clothing. Rising sea levels cut the island off from the Australian mainland in the early Holocene, resulting in a sharp drop in effective population size that

reduced the pool of social learners, resulting in these cultural losses (11).

Furthermore, because cultural transmission occurs vertically, from parents to children (12), dual inheritance theory considers cultural transmission as analogous to—but distinct from—genetic transmission (3). The use of methodologies better known for dealing with genetic data, such as cladistics and phylogenetic analysis, is beginning to yield valuable information on the rates, timings, and directions of such processes as the colonization of the Pacific Islands, the spread of agriculture across Europe from its Near eastern origins, and the changing compositions of pottery assemblages in the later stone age (13).

However, cultural transmission is always first and foremost social transmission, firmly embedded in networks of social relations between individuals. Thus, large-scale models can be enhanced by considering the small-scale processes revealed by ethnography. For example, among Micronesian sailors, the traditional skills of navigation are passed down between generations, often from father to son. From a very young age, children are immersed in discussion of canoes and navigation, and from the age of perhaps five upwards, teaching becomes more explicit. Knowledge is acquired through rote learning, the rehearsal of drills, chants, and stories, and the construction of “star charts” and “stick maps” that transmit the details of voyages covering more than 1400 miles of ocean (see the second photo). But this data is only part of the package; over more than 10 years, children are educated into a practical, physical understanding of how to use stars, ocean swells, currents, and wildlife in the actual performance of navigation (14).

In this case, the “maps” are deliberately designed to be ephemeral and thus leave no trace in the archaeological record. Nevertheless, the social expertise and relationships that underpin the transmission of navigational skills are fundamental to the negotiation and maintenance of wider social networks. These networks connect islands through practices such as gifting, trade, and exchange, leaving material traces that could be found in the archaeological record. Thus, the entanglement of cultural transmission with social relationships creates the patterns visible in the archaeological record. The adoption of social network models to investigate their interrelations is therefore an exciting new development in archaeology that has been used to model interactions of another island group—the

Aegean Cyclades—in the Bronze Age (15).

How and why human cultural transmission differs from that documented in other species is a fundamental question. As the only discipline with the temporal scope to investigate patterns and processes of cultural transmission from the first hominins to the modern day, archaeology is well placed to integrate the insights of the many disparate disciplines whose work informs on the question. We must learn to tack between the large scales of cultural transmission and the small scale of social relations to gain the best possible understanding of cultural transmission past and present.

References and Notes

1. W. C. McGrew, *Chimpanzee Material Culture: Implications for Human Evolution* (Cambridge Univ. Press, Cambridge, 1992).
2. C. P. van Schaik *et al.*, *Science* **299**, 102 (2003).
3. R. Boyd, P. J. Richerson, *The Origin and Evolution of Cultures* (Oxford Univ. Press, Oxford, 2005).
4. M. Tomasello, *Annu. Rev. Anthropol.* **28**, 509 (1999).
5. V. Gallese, in *Empathy and Fairness*, G. Bock, Jamie Goode, Eds. (Wiley, Chichester, 2006), pp. 3–19.
6. T. Matsuzawa, *Dev. Sci.* **10**, 97 (2007).
7. B. H. Smith, R. L. Tompkins, *Annu. Rev. Anthropol.* **24**, 257 (1995).
8. S. Semaw *et al.*, *Nature* **385**, 333 (1997).
9. D. Stout, T. Chaminade, *Neuropsychologia* **45**, 1091 (2007).
10. R. I. M. Dunbar, *Annu. Rev. Anthropol.* **32**, 163 (2003).
11. J. Henrich, *Am. Antiquity*, **69**, 197 (2004).
12. S. Shennan, J. Steele, in *Mammalian Social Learning: Comparative and Evolutionary Perspectives*, H. O. Box, K. R. Gibson, Eds. (Cambridge Univ. Press, Cambridge, 1999), pp. 367–388.
13. S. Shennan, *Genes, Memes and Human History: Darwinian Archaeology and Cultural Evolution* (Thames and Hudson, London, 2002).
14. S. D. Thomas, *The Last Navigator* (Hutchinson, London, 1987).
15. T. Evans, C. Knappett, R. Rivers, in *Complexity Perspectives on Innovation and Social Change*, D. Lane, D. Pumain, S. van der Leeuw, G. West, Eds. (Springer, Berlin, in press; final version available at www3.imperial.ac.uk/pls/portallive/docs/1/7292491.PDF).
16. The author acknowledges support from the British Academy Centenary Project Lucy to Language, the Archaeology of the Social Brain.

10.1126/science.1150898

NEUROSCIENCE

A Quiescent Working Memory

Stefano Fusi

When we perform a complex task like driving a car, we need to retain important information that will affect our behavior later on. For example, when we see a yield traffic sign, the image is not merely stored in our memory, but it is also “actively” held in mind so that we can react appropriately at the next crossroad. In such a case, we make use of what is known as “working memory” (1). It is widely believed that this type of memory is maintained by the persistent activity of a population of neurons. On page 1543 in this issue, Mongillo *et al.* (2) propose instead that the memory is stored in the efficacies of connections (synapses) among these neurons. This type of memory can be easily and rapidly reactivated after a period of neuronal quiescence.

Most previous models of working memory were inspired by experiments in which non-human primates were trained to hold in mind the identity or location of a sensory stimulus for a few seconds. During these periods of memory retention, sustained neuronal activity was observed in regions of the brain including the prefrontal cortex (3) and parietal cortex (4). The recorded activity was specific to the identity of a previously shown stimulus, suggesting that the memory of a stimulus might be stored in the pattern of persistent activity.

The mechanism for sustaining such neural activity likely involves the collective behavior

of a large number of interacting cells. Circuits of cortical neurons can be forged that sustain activity reverberations for times that greatly exceed the inherent time constant of every cell in the circuit (5, 6). Neurons generate strong electrical impulses (spikes) when they receive enough excitatory inputs from other connected neurons. These spikes cause a release of neurotransmitter molecules at the synaptic connections with other neurons, which triggers an electrical impulse in the postsynaptic cells. Neurons can excite each other so that each spike in one neuron causes an increasing number of other neurons to generate spikes, even in the absence of external stimulation. This growth of activity in a population of strongly interacting neurons increases until some regulatory mechanism stabilizes a constant average neuronal activity. In cortical neuronal circuits, coupling among neurons can be chosen so that there are several possible patterns of persistent activity (7), each corresponding to a different memory. An external stimulus simply “selects” one of the memories by activating the corresponding pattern of persistent activity.

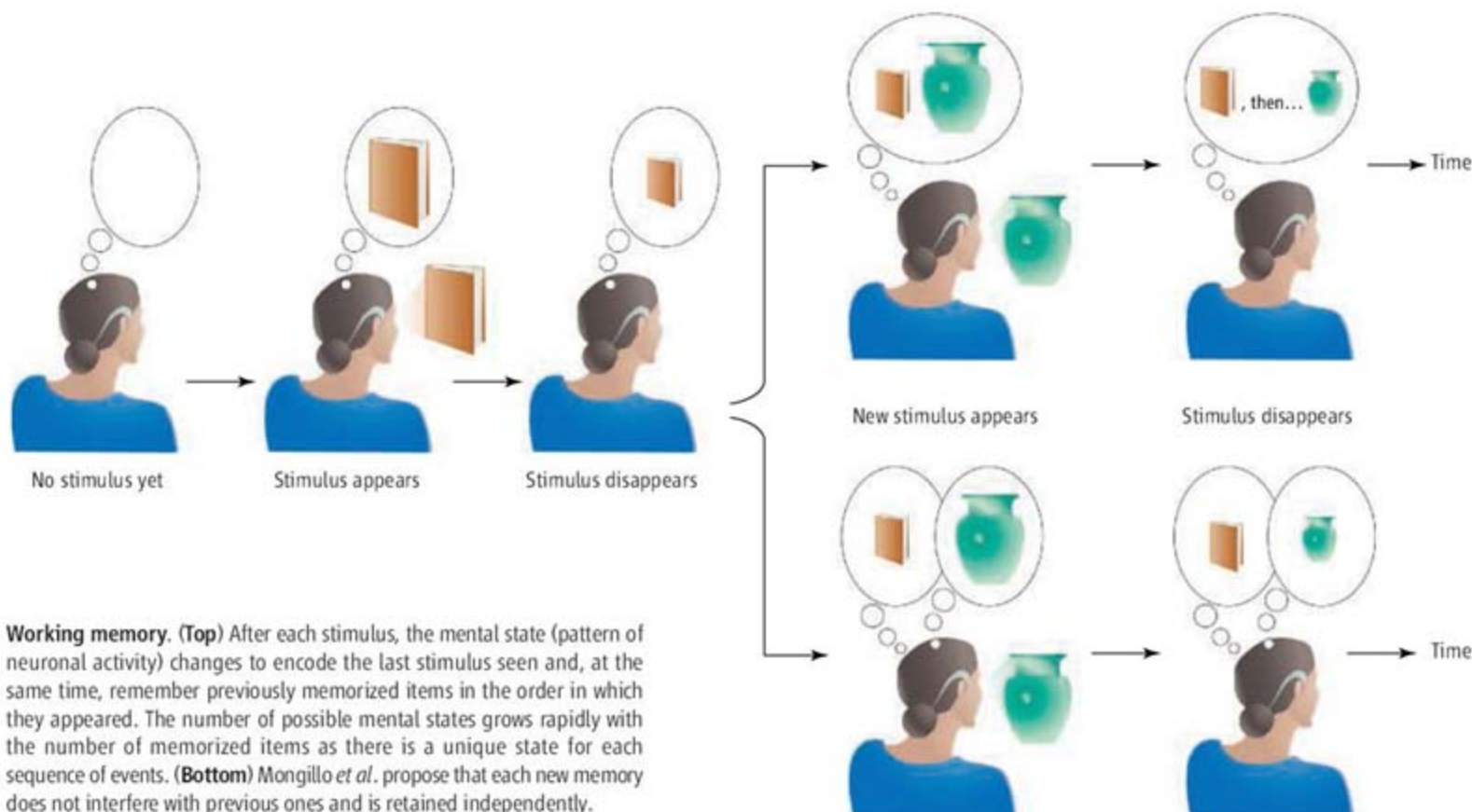
Mongillo *et al.* propose a new neural mechanism of working memory that is based on plasticity in synaptic connections rather than persistent neural activity. Thus, a memory resides in the pattern of synaptic strengths, and can be temporarily modified by a sensory stimulus to be remembered. Synaptic strengths are continuously modulated by spikes emitted by the presynaptic neurons (8). At each synapse, small-molecule neurotransmitters are released by the presynaptic neuron, and stimulate

Many of our actions or decisions are guided by what we experienced in the recent past.

receptors on the postsynaptic neuron. Each presynaptic spike not only depletes the neurotransmitter supply, but also increases the concentration of intracellular calcium. This, in turn, increases the amount of available neurotransmitters that can be released by the next presynaptic spike. Neurotransmitter depletion depresses synaptic strength until the resources are restored. However, the increased calcium concentration facilitates synaptic transmission and potentiates, temporarily, the synapse.

As with persistent activity, several possible memories are stored in the initial pattern of synaptic strengths, and a sensory stimulus selects a memory by activating one of the strongly interacting populations of neurons. The synaptic connections within this population are modulated by short-term depression and facilitation. After the stimulus is removed, total synaptic resources decrease due to depression, but the average fraction of resources used by each spike increases due to facilitation. The parameters can be chosen so that the net effect on the population activity is small, to the point that the spike activity is indistinguishable from the spontaneous activity preceding the stimulation. However, injection of a noisy current—such as when a memory recall signal is generated by another brain region—into a randomly selected subset of neurons is sufficient to reactivate the neuronal population that was previously stimulated. This indicates that memory of the sensory stimulus is still present. Indeed, any increase in neural activity facilitates all synapses. Those synapses originating from previously stimulated neurons are already

Center for Theoretical Neuroscience, Columbia University College of Physicians and Surgeons, New York, NY 10032, USA, and Institute of Neuroinformatics, ETH–University of Zurich, Zurich, 8057 Switzerland. E-mail: sf2237@columbia.edu



Working memory. (Top) After each stimulus, the mental state (pattern of neuronal activity) changes to encode the last stimulus seen and, at the same time, remember previously memorized items in the order in which they appeared. The number of possible mental states grows rapidly with the number of memorized items as there is a unique state for each sequence of events. (Bottom) Mongillo *et al.* propose that each new memory does not interfere with previous ones and is retained independently.

facilitated, and are strengthened even further. Eventually, each spike will activate an increasing number of neurons, leading to the activation of all cells of the previously stimulated population. After generating a spike, the neurons are quiet until the synapses recover from depression, and then a new “avalanche” will start. This oscillatory behavior will continue as long as the net synaptic efficacy remains above some critical level. Depending on the parameters of the input, the interplay between facilitation and depression can either sustain an oscillatory behavior, or a state of asynchronous enhanced activity.

This mechanism of Mongillo *et al.* has lower metabolic costs than persistent neural activity because it does not require the generation of spikes to retain memories. Moreover, it has two highly desirable features that complement those of the more traditional models of working memory. The first is that the memories are stored in the synaptic state and not in neural activity. This makes these memories quiescent until an external signal reactivates or refreshes them. During the silent period, the neural circuit responsible for holding the memory is decoupled from other brain areas because no spike is emitted, whereas any sustained activity would signal the participation of the cortical circuit in a larger network. Such a decoupling might play an important role in modular brain systems, in which different modules encode stimuli coming from differ-

ent sensory modalities. For example, cortical circuits located in different modules could independently store the memory of a large number of stimuli without interfering with other mental processes. These memories would be dynamically gated (9). When the memories are quiescent, the gate is closed, and information about the sensory stimuli is not communicated to other brain areas. When the memories are reactivated, the gate is open, and other neural circuits can rapidly acquire information contained in the stored memories. Gate opening might happen, for example, when attention is directed to a particular sensory modality, as in the case when you are reading a text (visual modality), and somebody asks whether the phone in the next room was ringing a few seconds ago (auditory modality).

The second attractive feature of the proposed model is the ability to memorize multiple items within the same neural circuit without requiring a reverberating stable state of neuronal activity for each combination of memory items. Earlier models require a number of reverberating activity states that grow exponentially with the number of memorized items (see the figure). Although building such a system is possible when the number of items is not too large, the mechanism proposed by Mongillo *et al.* is easier to implement. Indeed, in their case, because memories are stored in the synaptic states, a strong

interaction between different neurons is not required until the memories are retrieved. This allows us to hold in mind the identity of multiple stimuli that appear at different times without relying on a complex pattern of synaptic connections. Such a mechanism also permits us to retain memories in the face of distraction, as when having to remember one visual stimulus embedded in a long series of irrelevant distractors.

There may, nevertheless, be situations requiring a large number of stable states of reverberating activity—for example, complex tasks in which it is necessary to remember both the events and the temporal order in which they occur. In such a case, the interaction between different memories would play a fundamental role and a different mechanism might be required.

References

1. A. D. Baddeley, G. J. Hitch, in *The Psychology of Learning and Motivation: Advances in Research and Theory*, G. A. Bower, Ed. (Academic Press, New York, 1974), vol. 8, pp. 47–89.
2. G. Mongillo, O. Barak, M. V. Tsodyks, *Science* **319**, 1543 (2008).
3. J. M. Fuster, *J. Neurophysiol.* **36**, 61 (1973).
4. P. S. Goldman-Rakic, *Neuron* **14**, 477 (1995).
5. D. J. Amit, *Behav. Brain Sci.* **18**, 617 (1995).
6. X.-J. Wang, *Trends Neurosci.* **24**, 455 (2001).
7. J. J. Hopfield, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2554 (1982).
8. H. Markram, Y. Wang, M. Tsodyks, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5323 (1998).
9. R. C. O'Reilly, *Science* **314**, 91 (2006).

The Late Pleistocene Dispersal of Modern Humans in the Americas

Ted Goebel,^{1*} Michael R. Waters,² Dennis H. O'Rourke³

When did humans colonize the Americas? From where did they come and what routes did they take? These questions have gripped scientists for decades, but until recently answers have proven difficult to find. Current genetic evidence implies dispersal from a single Siberian population toward the Bering Land Bridge no earlier than about 30,000 years ago (and possibly after 22,000 years ago), then migration from Beringia to the Americas sometime after 16,500 years ago. The archaeological records of Siberia and Beringia generally support these findings, as do archaeological sites in North and South America dating to as early as 15,000 years ago. If this is the time of colonization, geological data from western Canada suggest that humans dispersed along the recently deglaciated Pacific coastline.

Explaining when and how early modern humans entered the New World and adapted to its varied environments is one of anthropology's most exciting and enduring questions. Until recently, it was generally believed that about 13.5 thousand years ago (ka) (1) the first migrants spread rapidly from Beringia to Tierra del Fuego in a few centuries, passing through an interior ice-free corridor in western Canada, becoming Clovis, and hunting to extinction the last of the New World's megamammals (2). Today, we realize that the peopling of the Americas was a much more complex process, because of two significant developments during the past decade. Molecular geneticists, using refined methods and an ever-increasing sample of living populations and ancient remains, are now capable of providing reliable information on the Old World origins of the first Americans, the timing of their initial migration to the New World, and the number of major dispersal events. Archaeologists have found new sites and reinvestigated old ones using new methods, to test whether a pre-13.5-ka population existed in North and South America, and to explain how early populations colonized its unpeopled landscapes (Fig. 1). Here, we review these developments and present a working model explaining the dispersal of modern humans across the New World. We focus primarily on molecular genetic, archaeological, and human skeletal evidence. We do not review the contributions of historical linguistics, because most linguists today are pessimistic about the use of their data

to reconstruct population histories beyond about 8 ka (3).

The Genetic Evidence

All human skeletal remains from the Americas are anatomically modern *Homo sapiens*; thus, the peopling of the New World is best understood in the context of the evolution and dispersal of modern humans in the Old World. Modern human dispersal from Africa across Eurasia began by about 50 ka (4, 5) and culminated with colonization of the Americas. Evidence from nuclear gene markers, mitochondrial (mt)DNA, and Y chromosomes indicates that all Native Americans came from Asia (6, 7). Molecular genetic diversity among modern Native Americans fits within five mtDNA (A, B, C, D, and X) and two Y-chromosome (C and Q) founding haplogroups, and all of these are found among indigenous populations of southern Siberia, from the Altai to Amur regions (8–10). Of these haplogroups, only X is known from both central Asia and Europe; however, X is a large, diverse haplogroup with many lineages, and the lineage found in Native American populations is distinct from those in Eurasia (6, 11). Ancient DNA from early American skeletal remains (12, 13) and human coprolites (14) link the present and the past; these, too, have only yielded Native American haplogroups of Asian origin. Based on the modern and ancient DNA records, then, Asia was the homeland of the first Americans, not Europe, lending no support to the recently proposed "Solutrean hypothesis," that the progenitors of Clovis were derived from an Upper Paleolithic population on the Iberian Peninsula (15, 16).

Using contemporary mtDNA and Y-chromosome variation as a clock, geneticists calculate that modern humans dispersed into greater central Asia by 40 ka (4, 17, 18), setting the stage for the colonization of the Americas. Corroborating human skeletal evidence of this event, however, is scarce. The earliest modern human remains in



Fig. 1. Map showing location of archaeological sites mentioned in text (1, Yamashita-cho; 2, Tianyuan Cave; 3, Studenoe-2; 4, Mal'ta; 5, Nizhniy Idzhir; 6, Alekseevsk; 7, Nepa-1; 8, Khaergas Cave; 9, Diuktai Cave; 10, Byzovaia; 11, Mamontovaya Kurya; 12, Yana RHS; 13, Ushki; 14, Tuluq; 15, Nogahabara I; 16, Nenana; 17, Swan Point; 18, Old Crow; 19, Bluefish Caves; 20, Kennewick; 21, Paisley Caves; 22, Spirit Cave; 23, Arlington Springs; 24, Calico; 25, Tule Spring; 26, Pendejo Cave; 27, La Sena and Lovewell; 28, Gault; 29, Schaefer, Hebior, and Mud Lake; 30, Meadowcroft Rockshelter; 31, Cactus Hill; 32, Topper; 33, Page-Ladson; 34, Tlapacoya; 35, Pedra Furada; 36, Lagoa Santa; 37, Pikimachay; 38, Quebrada Jaguay; 39, Quebrada Santa Julia; 40, Monte Verde; 41, Piedra Museo; 42, Cerro Tres Tatas and Cuevo Casa del Minero; 43, Fell's Cave).

Siberia are from Mal'ta and date to only 24 ka (19). In nearby eastern Asia, though, modern human fossils from Tianyuan Cave and Yamashita-cho are dated to the critical period, 39 to 36 ka (20), and in Siberia, archaeological evidence suggests that modern humans entered the region by 45 to 40 ka, when initial Upper Paleolithic technologies, tool forms, items of personal adornment, and art appeared for the first time (21). In Europe, archaeologists link the emergence of such behaviors to the spread of modern humans from southwestern Asia (22).

¹Center for the Study of the First Americans, Department of Anthropology, Texas A&M University, 4352 TAMU, College Station, TX 77843-4352, USA. ²Center for the Study of the First Americans, Departments of Anthropology and Geography, Texas A&M University, 4352 TAMU, College Station, TX 77843-4352, USA. ³Department of Anthropology, University of Utah, Salt Lake City, UT 84122-0060, USA.

*To whom correspondence should be addressed. E-mail: goebel@tamu.edu

Establishing when central Asian and Native American haplogroup lineages last shared a common ancestor has proven to be difficult. Current coalescent estimates based on variation in extant mtDNA lineages set the event at 25 to 20 ka (4) or less than 20 ka (23), after the last glacial maximum (LGM), and estimates based on Y-chromosome variability suggest that divergence occurred after 22.5 ka, possibly as late as 20 to 15 ka (7, 24, 25). The differences in calculations are the result of several issues, including potential variation in mutation rates, variable and sometimes circular techniques of calibrating coalescent times to calendar years, time-dependency of mutation and/or substitution rates, and effects of genetic drift on the original founding population (4, 26).

New analyses of haplogroup subclades help to resolve when modern humans subsequently spread from Beringia to the rest of the Americas. Three subclades of mtDNA subhaplogroup C1 are widely distributed among North, Central, and South Americans but absent in Asian populations, which suggests that they evolved after the central Asian–Native American split, as the first Americans were dispersing from Beringia (27). The estimated date of coalescence for these subclades is 16.6 to 11.2 ka, which suggests that the colonization of the Americas south of the continental ice sheets may have occurred sometime during the late-glacial period, thousands of years after the initial splitting of Asian and Native American lineages. Genetic simulation studies and analyses of the geographic structure of Native American mtDNA haplogroups further suggest that colonization from Beringia occurred earlier in this time frame (about 16 ka) than later, because late-entry, rapid-spread models (like the Clovis-First model) are not capable of generating the observed geographic distribution of genetic patterns in extant populations (28, 29).

The cranial morphology of the earliest Americans [i.e., “Paleoamericans” like Kennewick (Washington), Spirit Cave (Nevada), and Lagoa Santa (Brazil)] is significantly different from that of more recent Native Americans (30). Given the assumption that craniometric variation is neutral and therefore phylogenetically significant, the differences could reflect two successive migrations stemming from two geographically or temporally distinct sources (31–33). Accordingly, Paleoamericans came to the New World first and were later replaced by ancestors of modern Native Americans.

Genetic data do not support this model. All major Native American mtDNA and Y-chromosome haplogroups emerged in the same region of central Asia, and all share similar coalescent dates, indicating that a single ancient gene pool is ancestral to all Native American populations (6, 10, 16). Similarly, all sampled native New World populations (from Alaska to Brazil) share a unique allele at a specific microsatellite locus that is not found in any Old World populations (except Koryak and Chukchi of western

Beringia), which implies that all modern Native Americans descended from a single source population (34). This history is further supported by ancient DNA studies showing that Paleoamericans carried the same haplogroups (and even subhaplogroups) as modern Native groups (12–14). Thus, although the Paleoamerican sample is still small, the morphological differences are likely the result of genetic drift and natural selection (30), not separate migrations.

A separate but related problem is whether some modern Native American populations resulted from migrations that occurred after initial human dispersal. Phylogenetic analyses of haplogroup lineages cannot easily discriminate between a single migration and multiple migrations of genetically distinct but closely related populations. For this, we need identification of specific mtDNA and Y-chromosome haplogroup subclades through analysis of the entire molecule (as well as detailed studies of nuclear genome variation). A recent study investigating mtDNA subclade distributions across Siberia (11) recognized two subclades of haplogroup D2, one among central Siberian groups (D2a) and the other among Chukchi, Siberian Eskimos, and Aleuts (D2b). These subclades share a coalescent date of 8 to 6 ka, which suggests that middle-Holocene ancestors of modern Eskimo-Aleuts spread from Siberia into the Bering Sea region and not vice versa, which supports earlier interpretations based on dental evidence (35).

The Archaeological Evidence

To colonize the Americas, modern humans had to learn to subsist in the extreme environments of the Siberian Arctic. They did this by 32 ka. The evidence comes from the Yana Rhinoceros Horn site (RHS), which is located along the lower Yana River in northwest Beringia and contains a frozen, well-preserved cultural layer with stone artifacts and remains of extinct fauna (36). Most interesting are bi-beveled rods on rhinoceros horn and mammoth ivory, signs of a sophisticated Upper Paleolithic technology. Sites of similar age occur in subarctic central Siberia (Nepa, Alekseevsk) and arctic European Russia (Mamontovaya Kurya, Byzovaia) (21, 37), which suggests that people had become well-equipped to handle life in the far north shortly after arriving in south Siberia (22). Their spread into the Arctic occurred during a time of relatively warm climate before the LGM.

As yet, no unequivocal traces that the early people of Yana RHS explored farther east onto the Bering Land Bridge and crossed into Alaska and northwest Canada have been found, but hints of an early human presence may include the 28-ka mammoth-bone core and flake recovered from Bluefish Caves (Yukon Territory) and even older bone materials from along the nearby Old Crow River (38). These bones, however, lack associated stone artifacts and might be the result of natural bone breakage (39). Instead, the earliest reliable archaeological evidence from

eastern Beringia comes from Swan Point (central Alaska), where a distinctive microblade and burin industry dates to 14 ka (40). The Swan Point artifacts share many technological qualities with late Upper Paleolithic sites in central Siberia (e.g., Studenoe-2, Nizhnii Idzhir, Khaergas Cave, Diuktai Cave) and appear to document the dispersal of microblade-producing humans from Siberia to Beringia during the late glacial.

After 14 ka, the Beringian archaeological record becomes much more complicated. The best-documented industries for this time are the Nenana complex of central Alaska (dating to 13.8 to 13 ka) and the early Ushki complex of Kamchatka (13 ka) (22, 41). These complexes contain small bifaces and unifaces made on blades and flakes, but they lack microblades and burins. The Sluiceway-Tuluq complex (northwest Alaska) also may be contemporaneous to Nenana but is technologically distinct from it (22). These sites contain large lanceolate bifaces that appear to date to about 13.2 ka. Another site, Nogahabara I (west Alaska), contains a mixed array of artifacts (lanceolate bifaces, notched bifaces, and microblade cores) reportedly dated to 13.8 to 12.7 ka (42); however, this site must be viewed with caution because the artifacts and bones used for dating are from near-surface and surface contexts in a sand dune blowout, a context notorious for artifact redeposition and mixing. After 13 ka, microblade and burin technologies reappear, sometimes in combination with bifacial point technologies. Perhaps these changes through time and across space relate to cultural differences and population turnovers, but more likely they represent the development of a unique human adaptation to the rapidly changing shrub-tundra environment of late-glacial Beringia (22). A small number of undated fluted points similar to Clovis occur in Alaska (39), but their relation to Clovis points found south of the continental ice sheets is unknown and may represent the backward flow of technologies (or people) from mid-latitude North America to Beringia at the very end of the Pleistocene (22, 39).

Since 40 ka, the Cordilleran and Laurentide ice sheets covered much of Canada, but during warmer periods they retreated sufficiently to create ice-free corridors along the Pacific coast and Plains east of the Canadian Rockies. These corridors were the conduits through which the first humans spread from Beringia to the Americas. When humans arrived in arctic Siberia at Yana RHS 32 ka, contracted ice sheets left wide-open corridors through which humans could have passed, but by 24 ka the ice sheets had grown sufficiently to clog both passageways (43). Although isolated ice-free refugia probably existed in both corridors throughout the LGM, humans probably did not occupy these areas until the corridors reopened during the late glacial. Timing of the reopening of the coastal and interior corridors is still debated, because of imprecise dating and because the various

Cordilleran glaciers reacted differently to climate change (43). Nonetheless, the coastal corridor appears to have become deglaciated and open to human habitation by at least 15 ka, whereas the interior corridor may not have opened until 14 to 13.5 ka (44, 45). The archaeological records of both corridors are still inadequate for addressing questions about the initial peopling of the Americas; however, the presence of human remains dating to 13.1 to 13 ka at Arlington Springs, on Santa Rosa Island off the coast of California, indicates that the first Americans used watercraft (46).

Clovis and its contemporaries. Discussion of the early archaeological record south of the Canadian ice sheets starts with Clovis, the best-documented early complex in the Americas. Radiocarbon dates obtained over the last 40 years from Clovis sites across North America suggested that the complex ranged in age from 13.6 to 13 ka (2); however, evaluation of the existing dates and new ^{14}C assays reveals that Clovis more precisely dates from 13.2–13.1 to 12.9–12.8 ka (47), a shorter and younger time span for Clovis than earlier thought. The current evidence suggests Clovis flourished during the late Allerød interstadial and quickly disappeared at the start of the Younger Dryas stadial. The apparent simultaneous appearance of Clovis across much of North America suggests that it rapidly expanded across the continent, but the overlap in ^{14}C dates between regions of North America makes it impossible to determine a point of origin or direction of movement.

With recently excavated Clovis assemblages, especially from the southeastern United States and Texas, we know unequivocally that Clovis is characterized by not only bifacial technology but also distinctive Upper Paleolithic blade technology (Fig. 2) (15, 48). The principal diagnostic artifact of Clovis is its lanceolate fluted projectile point, not just because of its form but also the technology used to produce it. Other tool forms were equally important, especially formal stone tools like end scrapers, as well as cylindrical rods made on ivory, antler, or bone. These rods were beveled at one or both ends and functioned as fore-shafts or projectile points, respectively (48).

Traditionally, Clovis has been thought to represent a population of mobile hunter-gatherers because individual Clovis tools had multiple functions and were highly curated, which suggests that they were part of a conveniently transported tool kit (2). Many Clovis tools were made on high-quality stones like chert and obsidian procured hundreds of kilometers from where they were eventually discarded (48). Clovis sites are small and typically represent mammoth or mastodon kills, short-term camps, or caches. In the southeastern United States and Texas, however, enormous scatters of Clovis artifacts have been found that possibly represent quarry-habitation sites habitually used by Clovis people, from which they did not range great distances.

At the Gault site (Texas), of 650,000 excavated artifacts (mostly debitage), 99% are made from local, on-site cherts; rare nonlocal materials are from sources only 70 km away (49).

Clovis points have long been known to be associated with remains of mammoth and mastodon (2), but the importance of proboscideans in Clovis subsistence remains uncertain. Optimal foraging theory has been used to predict that humans would not become proboscidean-hunting specialists (50), and certainly the recurrence of bison, deer, hares, reptiles, and amphibians indicates that, in some contexts, Clovis people did more than hunt mammoth and mastodon (51). However, at least 12 unequivocal Clovis proboscidean kill and butchery sites are known (52), an unusually high number for such a short period of time, given that there are only six proboscidean kill sites for the entire Eurasian Upper Paleolithic (53). In most areas of North America, Clovis people hunted mammoth and mastodon regularly, and they likely played some

role in their extinction. It is not surprising that they also subsisted on a variety of other foods.

Most Clovis sites are in North America. Few Clovis artifacts have been found in Central and South America (54). Instead, a different complex of archaeological sites may mark this era south of Panama. At least six sites in South America (Cerro Tres Tetras, Cueva Casa del Minero, and Piedra Museo in Argentina and Fell's Cave, Quebrada Santa Julia, and Quebrada Jaguay in Chile) have multiple dates that overlap the known age of Clovis (47, 55, 56). These sites mostly contain undiagnostic flake tools and bifaces, but distinctive Fishtail points (some with fluted bases) were found in deposits dating to 13.1 to 12.9 ka at Fell's Cave and Piedra Museo. Although it has been suggested that Fishtail points postdate Clovis and were derived from it (54), the two may have shared an earlier, as yet unidentified progenitor. Among the newest Clovis-aged localities in South America is Quebrada Santa Julia, a stratified site with a well-preserved

living floor and hearth dating to 13.1 ka (57). Associated with the hearth were a broken, nondiagnostic, fluted biface, several flake tools, a core, and nearly 200 flakes, as well as remains of extinct horse. Quebrada Santa Julia provides an unambiguous association of fluting technology and extinct fauna in South America.

Early occupations. Since the discovery and definition of Clovis, researchers have searched for evidence of an even older occupation of the Americas, but most sites dating before Clovis investigated between 1960 and 1995 [e.g., Calico (California), Tule Springs (Nevada), Pendejo Cave (New Mexico), Pedra Furada (Brazil), Pikimachay Cave (Peru), and Tlapacoya (Mexico)] have not held up to scientific scrutiny (2, 39). Perhaps the best candidate is the Monte Verde site (Chile), which contains clear artifacts in a sealed context and is dated to 14.6 ka (58). Despite criticism (59), its acceptance by most archaeologists means synchronous and possibly earlier sites should exist in North America. A few localities dating between 15 and 14 ka now seem to provide compelling evidence of an occupation before Clovis.

In the northern United States, the Schaefer and Hebior sites (Wisconsin) provide strong evidence of human proboscidean hunting or scavenging near the margin of the Laurentide ice sheet between 14.8 and 14.2 ka (60, 61). At each site, disarticulated remains of a single mammoth were sealed in pond clay and associated with unequivocal

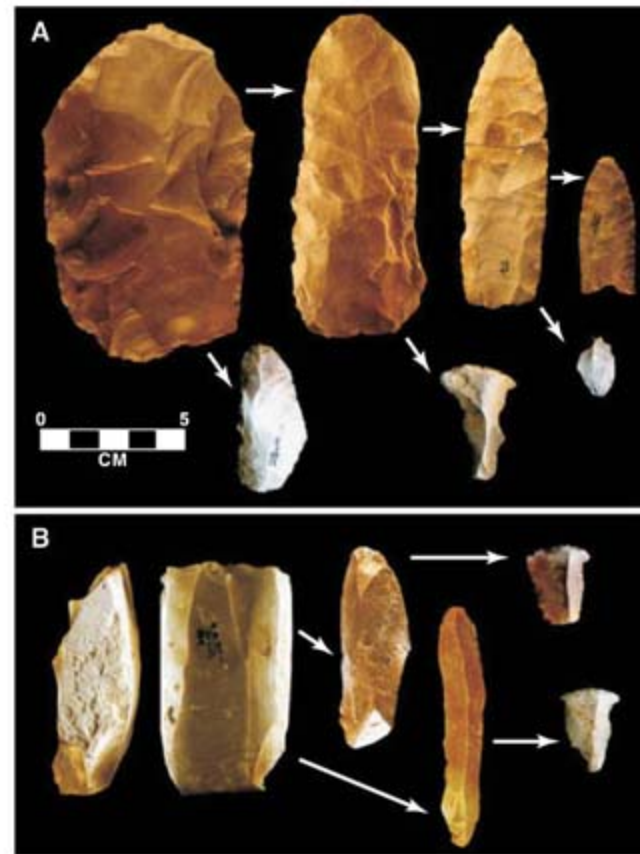


Fig. 2. The distinctive Clovis biface and blade technologies (schematic diagram with approximate scale). Clovis fluted points were manufactured by reduction of a large blank through a succession of stages including removal of broad thinning flakes across the entire face of the biface, end thinning at all stages, and final fluting of the finished piece (A). Thinning flakes were often utilized as tools. Clovis blades were detached from conical and wedge-shaped cores (B), the main distinction being that conical cores have blade removals around their entire circumference while wedge-shaped cores have a single front of blade removals. Blades are long, parallel-sided, curved in longitudinal cross section, and triangular or trapezoidal in transverse cross section; they were often used as tools. These specific artifacts are made on Edwards chert from the Gault site, Texas.

stone artifacts. The bones bear consistent signs of butchering—cut and pry marks made by stone tools (61). Critics suggest that the bone breakage and surface marring is the result of natural processes (2); however, it is difficult to reject the evidence from these sites because of the consistent patterning of the marks, low-energy depositional context, and associated stone tools. Even earlier evidence of humans in Wisconsin is suggested by what appear to be cut and pry marks on the lower limb bones of a mammoth recovered from Mud Lake. These bones date to 16 ka, but stone tools are absent (61).

Three other sites—Meadowcroft Rockshelter (Pennsylvania), Page-Ladson (Florida), and Paisley Cave (Oregon)—may provide additional evidence of humans in North America by about 14.6 ka. At Meadowcroft Rockshelter, artifacts occur in sediments that may be as old as 22 to 18 ka (62), but it is the record post-dating 15.2 ka that is especially interesting. This is the uppermost layer of lower stratum IIa, which produced a small lanceolate biface and is bracketed by dates of 15.2 and 13.4 ka. Acceptance of the site, however, hinges on resolution of dating issues (63).

At Page-Ladson, early materials occur in a buried geologic context within a sinkhole that is now submerged by the Aucilla River. Seven pieces of chert debitage, one expedient unifacial flake tool, and a possible hammerstone were associated with extinct faunal remains, including a mastodon tusk with six deep grooves at the point where the tusk emerged from the alveolus of the cranium (64). These grooves are interpreted to have been made by humans as the tusk was removed from its socket. Seven ^{14}C dates for this horizon average about 14.4 ka, which suggests human occupation of the sinkhole during the late Pleistocene when the water table was lower than it is today. Page-Ladson may contain evidence of pre-Clovis humans, but, despite extensive reporting on the site, more details on artifact contexts and site formation processes are needed to permit objective evaluation of the record.

At Paisley Cave, three human coprolites are directly ^{14}C dated to about 14.1 ka (14). The human origin of the coprolites is supported by ancient mtDNA analyses that showed they contained haplogroups A and B, but a complete report is not yet available.

The evidence for humans in the Americas even earlier than 15 ka is less secure, but recently has been presented for four sites: Cactus Hill (Virginia), La Sena (Nebraska), Lovewell (Kansas), and Topper (South Carolina). Cactus Hill is a sand-dune site with late prehistoric,

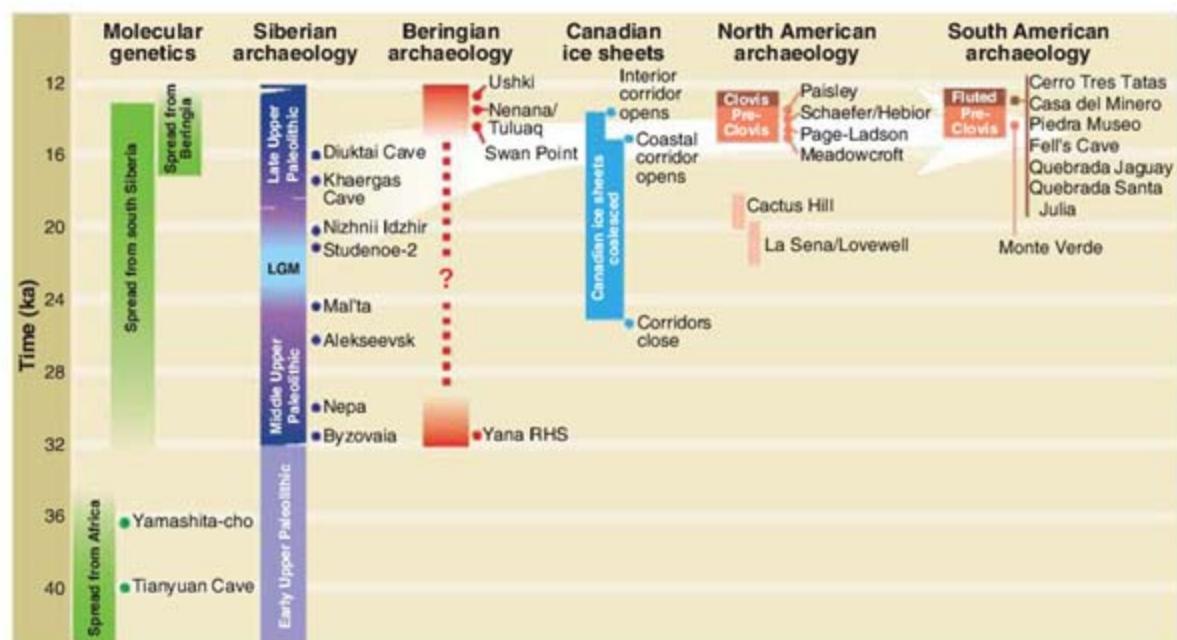


Fig. 3. Combined, the molecular genetic and archaeological records from Siberia, Beringia, and North and South America suggest humans dispersed from southern Siberia shortly after the last glacial maximum (LGM), arriving in the Americas as the Canadian ice sheets receded and the Pacific coastal corridor opened, 15 ka.

Archaic, and Clovis levels. Potentially older artifacts, including small prismatic blade cores, blades, and two basally thinned bifacial points were recovered 10 to 15 cm below the Clovis level (65). Three ^{14}C dates ranging from 20 to 18 ka are reported from the levels below Clovis, but there are also dates of 10.3 ka and later. Charcoal samples were not recovered from hearth features but occur as isolated fragments at the same level as the artifacts. The younger dates indicate translocation of charcoal from overlying sediments, and the older charcoal could be derived from sediments underlying the cultural layer (59, 63), but luminescence dates on the aeolian sands correlate with the older ^{14}C results and indicate minimal mixing of the sediments (66). Even though much information has yet to be published about this site, the potential presence of a biface and blade assemblage stratigraphically below the site's Clovis assemblage is compelling.

An even older occupation has been proposed based on taphonomically altered mammoth bones at the La Sena and Lovewell sites that date from 22 to 19 ka (67). Neither site has yielded stone tools or evidence of butchering; however, many of the leg bones display percussion impact and flaking, which suggests that they were quarried and flaked by humans while they were in a fresh, green state, within a few years of the death of the animals. Clovis people periodically flaked bone in this fashion, as did Upper Paleolithic Beringians (2, 22); however, in those contexts humans left behind stone tools, whereas at La Sena and Lovewell stone tools are absent.

Currently, the oldest claim for occupation of North America is at the Topper site, located on a Pleistocene terrace overlooking the Savannah River. Clovis artifacts at Topper are found at the

base of a colluvial deposit, and older artifacts are reported in underlying sandy alluvial sediments dated to about 15 ka (68). The proposed early assemblage is a smashed core and microlithic industry. Cores and their removals show no negative bulbs, and flakes and spalls were modified into small unifacial tools and "bend-break tools," possibly used for working wood or bone. In 2004, similar-looking material was found in older alluvial deposits dating in excess of 50 ka (69). Given that the assemblage was not produced through conventional Paleolithic technologies and that the putative artifacts could have been produced through natural processes (specifically thermal spalling), evaluation of this site must await a complete lithic analysis.

Combining the Genetic and Archaeological Evidence

Unquestionably, the human skeletal evidence across the Americas shows that the New World was populated by *Homo sapiens*. Although the crania of these early people look different from modern Native Americans, modern and ancient DNA studies show that they were genetically related. The earliest inhabitants of the Americas hailed from south Siberia (between the Altai Mountains and Amur valley) and ultimately descended from a population of modern humans who dispersed from Africa by 50 ka and appeared in central Asia by 40 ka. Thus, a maximum limiting age can be placed on the entry of people into the New World of no earlier than 40 ka. Any claims for an earlier migration should be viewed with skepticism.

Current molecular evidence implies that members of a single population left Siberia and headed east to the Americas sometime between about 30 and 13 ka (Fig. 3). Most studies suggest this event occurred after the LGM, less than 22 ka. Recent

analyses of mtDNA and nuclear sequence data further suggest a dispersal south from Beringia after 16.6 ka (27), from a founding population of less than 5000 individuals (70). The genetic record has not revealed multiple late-Pleistocene migrations, but does distinguish a Holocene dispersal of Eskimo-Aleuts from northeast Asia. There is nothing in the modern or ancient genetic records to suggest a European origin for some Native Americans.

At first glance, the genetic evidence would seem to mesh well with the traditional view that Clovis represents the first people to enter the Americas. Redating of Clovis from 13.2–13.1 to 12.9–12.8 ka indicates it is not only centuries younger than the late-glacial complexes of Alaska but also younger than even the most conservative estimate for the opening of the interior Canadian corridor. The Clovis-First model, however, requires all American sites older than Clovis to be rejected, and this appears to be no longer possible. The Clovis-First model does not explain the apparent synchronicity between Clovis and the early Paleo-Indian sites of South America. Finally, a late-entry and rapid dispersal of humans across the New World is inconsistent with the distribution of genetic variation observed in Native American populations today.

Humans possibly colonized the Americas before the LGM. They occupied western Beringia by 32 ka, and no glacial ice sheets would have blocked passage through western Canada during this relatively warm time. However, there is still no unequivocal archaeological evidence in the Americas to support such an early entry.

The most parsimonious explanation of the available genetic, archaeological, and environmental evidence is that humans colonized the Americas around 15 ka, immediately after deglaciation of the Pacific coastal corridor. Monte Verde, Schaefer, and Hebior point to a human presence in the Americas by 14.6 ka. Human occupations at Meadowcroft, Page-Ladson, and Paisley Cave also appear to date to this time. Together these sites may represent the new basal stratum of American prehistory, one that could have given rise to Clovis. Most mtDNA and Y-chromosome haplogroup coalescence estimates predict a 15-ka migration event, and it may correlate to the post-LGM dispersal of microblade-producing populations into northern Siberia and their eventual appearance in Beringia during the late glacial. The first Americans used boats, and the coastal corridor would have been the likely route of passage since the interior corridor appears to have remained closed for at least another 1000 years. Once humans reached the Pacific Northwest, they could have continued their spread southward along the coast to Chile, as well as eastward along the southern margin of the continental ice sheets, possibly following traces of mammoth and mastodon to Wisconsin. Clovis could have originated south of the continental ice sheets, and the dense Clovis quarry-campsites in the southeastern United States may

be the result of a longer occupation there than in other regions. Alternatively, Clovis could be the result of a second dispersal event from Beringia to America—from the same ancestral gene pool as the first dispersing population—when the interior ice-free corridor opened, about 13.5 ka.

The peopling of the Americas debate is far from resolved. To move forward, we must continue to take an interdisciplinary scientific approach to the problem. Archaeological investigations will provide the empirical evidence of the first Americans, but this evidence must be objectively and rigorously evaluated. Geoarchaeological investigations have and will play a major role by documenting the geological and geochronological context of sites and developing predictive models to find early sites. The sparse evidence for pre-13 ka occupation of the Americas may be a problem of sampling and artifact recognition. Genetic studies will also be key as more is learned about modern and ancient haplogroup subclades in combination with full mtDNA genome sequencing and identification of patterns of nuclear DNA variation. The empirical data from these fields and other disciplines will ultimately provide the evidence needed to build and test models to explain the origins and dispersal of the first Americans.

References and Notes

- All ages are presented as ka (thousands of calendar years ago). Dates relating to genetic events are in calendar years based on coalescent methods. Dates relating to archaeological events are derived by calibrating radiocarbon ages. Radiocarbon dates younger than 21,000 ^{14}C years ago were calibrated with Calib 5.0.1 (IntCal04 curve); older dates were calibrated by using CalPal Online (CalPal 2007 HULU curve).
- G. A. Haynes, *The Early Settlement of North America: The Clovis Era* (Cambridge Univ. Press, Cambridge, 2002).
- J. N. Hill, in *The Settlement of the American Continents: A Multidisciplinary Approach to Human Biogeography*, C. M. Barton, G. A. Clark, D. R. Yesner, G. A. Pearson, Eds. (Univ. of Arizona Press, Tucson, 2004), pp. 39–48.
- P. Forster, *Philos. Trans. R. Soc. London B Biol. Sci.* **359**, 255 (2004).
- M. Metspalu, T. Kivisild, H.-J. Bandelt, M. Richards, R. Villems, *Nucleic Acids Mol. Biol.* **81**, 181 (2006).
- D. A. Merriwether, in *Environment, Origins, and Population*, D. H. Ubelaker, Ed., *Handbook of North American Indians*, vol. 3, W. C. Sturtevant, Ed. (Smithsonian Institution Press, Washington, DC, 2006), pp. 817–830.
- T. M. Karafet, S. L. Zegura, M. F. Hammer, in *Environment, Origins, and Population*, D. H. Ubelaker, Ed., *Handbook of North American Indians*, vol. 3, W. C. Sturtevant, Ed. (Smithsonian Institution Press, Washington, DC, 2006), pp. 831–839.
- M. V. Derenko et al., *Am. J. Hum. Genet.* **69**, 237 (2001).
- E. B. Starikovskaya et al., *Ann. Hum. Genet.* **69**, 67 (2005).
- S. L. Zegura, T. M. Karafet, L. A. Zhivotovsky, M. F. Hammer, *Mol. Biol. Evol.* **21**, 164 (2004).
- M. Derenko et al., *Am. J. Hum. Genet.* **81**, 1025 (2007).
- B. M. Kemp et al., *Am. J. Phys. Anthropol.* **132**, 605 (2007).
- D. G. Smith, R. S. Malhi, J. A. Eshleman, F. A. Kaestle, B. M. Kemp, in *Paleoamerican Origins: Beyond Clovis*, R. Bonnichsen, B. T. Lepper, D. Stanford, M. R. Waters, Eds. (Center for the Study of the First Americans and Texas A&M Univ. Press, College Station, TX, 2005), pp. 243–254.
- D. L. Jenkins, in *Paleoindian or Paleoarchaic? Great Basin Human Ecology at the Pleistocene/Holocene*

- Transition, K. E. Graf, D. N. Schmitt, Eds. (Univ. of Utah Press, Salt Lake City, 2007), pp. 57–81.
- B. Bradley, D. Stanford, *World Archaeol.* **36**, 459 (2004).
- S. Wang et al., *PLoS Genet.* **3**(11), e185 (2007).
- R. S. Wells et al., *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10244 (2001).
- D. Comas et al., *Eur. J. Hum. Genet.* **12**, 495 (2004).
- M. P. Richards, P. B. Pettit, M. C. Stiner, E. Trinkaus, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6528 (2001).
- H. Shang, H. Tong, S. Zhang, F. Chen, E. Trinkaus, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 6573 (2007).
- T. Goebel, *Evol. Anthropol.* **8**, 208 (1999).
- J. F. Hoffecker, S. A. Elias, *Human Ecology of Beringia* (Columbia Univ. Press, New York, 2007).
- T. G. Schurr, S. T. Sherry, *Am. J. Hum. Biol.* **16**, 420 (2004).
- M.-C. Bortolini et al., *Am. J. Hum. Genet.* **73**, 524 (2003).
- M. Seielstad et al., *Am. J. Hum. Genet.* **73**, 700 (2003).
- S. Y. W. Ho, G. Larson, *Trends Genet.* **22**, 79 (2006).
- E. Tamm et al., *PLoS ONE* **2**(9), e829 (2007).
- A. G. Fix, *Am. J. Phys. Anthropol.* **128**, 430 (2005).
- D. H. O'Rourke, M. G. Hayes, S. W. Carlyle, *Hum. Biol.* **72**, 15 (2000).
- J. F. Powell, *The First Americans: Race, Evolution, and the Origin of Native Americans* (Cambridge Univ. Press, Cambridge, 2005).
- R. L. Jantz, D. W. Owsley, *Am. J. Phys. Anthropol.* **114**, 146 (2001).
- R. González-José et al., *Am. J. Phys. Anthropol.* **128**, 772 (2005).
- W. A. Neves, M. Hubbe, L. B. Piló, *J. Hum. Evol.* **52**, 16 (2007).
- K. B. Schroeder et al., *Biol. Lett.* **10.1098/rsbl.2006.0609** (2007).
- J. H. Greenberg, C. G. Turner II, S. L. Zegura, *Curr. Anthropol.* **27**, 477 (1986).
- V. V. Pitulko et al., *Science* **303**, 52 (2004).
- P. Pavlov, J. I. Svendsen, S. Indrelid, *Nature* **413**, 64 (2001).
- R. E. Morlan, *Quat. Res.* **60**, 123 (2003).
- E. J. Dixon, *Bones, Boats, and Bison: Archeology and the First Colonization of Western North America* (Univ. of Utah Press, Salt Lake City, 1999).
- C. E. Holmes, B. A. Crass, paper presented at the 30th annual meeting of the Alaska Anthropological Association, Fairbanks, 27 to 29 March 2003.
- T. Goebel, M. R. Waters, M. A. Dikova, *Science* **301**, 501 (2003).
- D. Odess, J. T. Rasic, in *American Antiquity* **72**, 691 (2007).
- J. J. Clague, R. W. Mathewes, T. A. Ager, in *Entering America: Northeast Asia and Beringia before the Last Glacial Maximum*, D. B. Madsen, Ed. (Univ. of Utah Press, Salt Lake City, 2004), pp. 63–94.
- C. A. S. Mandryk, H. Josenhans, D. W. Fedje, R. W. Mathewes, *Quat. Sci. Rev.* **20**, 301 (2001).
- A. S. Dyke, in *Quaternary Glaciations—Extent and Chronology, Part II: North America*, J. Ehlers, P. L. Gibbard, Eds. (Elsevier, Amsterdam, 2004), pp. 373–424.
- J. R. Johnson, T. W. Stafford Jr., G. J. West, T. K. Rockwell, American Geophysical Union Joint Assembly, Acapulco, 22 to 25 May 2007, *Eos* **88**(23), Jt. Assem. Suppl., Abstr. PP42A-03.
- M. R. Waters, T. W. Stafford Jr., *Science* **315**, 1122 (2007).
- K. B. Tankersley, in *The Settlement of the American Continents: A Multidisciplinary Approach to Human Biogeography*, C. M. Barton, G. A. Clark, D. R. Yesner, G. A. Pearson, Eds. (Univ. of Arizona Press, Tucson, 2004), pp. 49–63.
- M. B. Collins, in *Foragers of the Terminal Pleistocene in North America*, R. B. Walker, B. N. Driskell, Eds. (Univ. of Nebraska Press, Lincoln, 2007), pp. 59–87.
- D. A. Byers, A. Ugan, *J. Archaeol. Sci.* **32**, 1624 (2005).
- M. D. Gannon, D. J. Meltzer, *Quat. Sci. Rev.* **23**, 1955 (2004).
- D. K. Grayson, D. J. Meltzer, *J. Archaeol. Sci.* **30**, 585 (2003).
- T. Surovell, N. Waguespack, P. J. Brantingham, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 6231 (2005).
- J. E. Morrow, C. Gnecco, Eds., *Paleoindian Archaeology: A Hemispheric Perspective* (Univ. Press of Florida, Gainesville, 2006).
- L. Miotti, M. C. Salemme, *Quat. Int.* **109-110**, 95 (2003).

Predators Induce Cloning in Echinoderm Larvae

Dawn Vaughn* and Richard R. Strathmann

Asexual propagation (cloning) is a widespread reproductive strategy of plants and adult and larval animals. Possible advantages of larval cloning include increased fecundity, dispersal, and salvaging of material otherwise lost at metamorphosis (1–4). These advantages require a larval environment conducive to growth and reproduction. In contrast, we show here that echinoderm larvae clone in response to cues from fish predators, that is, under conditions indicating potential risk to larvae. Our results suggest that asexual reproduction can be a strategy for rapid size reduction as a defense.

We exposed 4-day-old sand dollar (*Dendraster excentricus*) larvae (plutei) to external mucus from fish (predator cues) (5). Within 24 hours, clones formed as buds aborally (Fig. 1A) or by anterior-posterior fission of early-stage larvae. Cloning occurred in all treatments with fish mucus: 20% at 0.01 mg/ml of mucus/ml of seawater, 40% at 0.1 mg/ml, 10% at 1.0 mg/ml. Cloning was not observed in controls without fish mucus. Buds detached as evenly ciliated gastrulae and developed into small larvae. All products of

cloning by budding and fission were smaller than uncloned larvae.

The response to fish mucus suggests that larval cloning may counter a threat of predation. Slow-swimming plutei have no effective escape from predatory fish; cloning may provide a temporary refuge from predation by producing two smaller, less-detectable propagules, a gastrula-like bud and a smaller-than-usual pluteus. Size reduction by cloning is slow compared with a fish attack, but the stimulus is presence of fish, as indicated by mucus, not an immediate attack by a fish.

Small size can protect zooplankton against predators. Smaller size at a given stage reduces vulnerability of fish larvae to visual predators (6). Small sand dollar larvae (from half-embryos) are less vulnerable to a visual predator (postlarval fish) although more susceptible to some nonvisual predators (7). However, there is evidence that small size in some zooplankton decreases vulnerability to nonvisual predators (8). We hypothesize that reduced size in clones decreases detection and selection by visual predators and may reduce signals to some nonvisual planktonic predators.

Reduced larval size resulting from predator-induced cloning also challenges the common assumption that bigger offspring are safer offspring. Cloning in plutei occurs in response to environmental information unavailable to parents on the sea floor. Adult sand dollars cannot assess risks to their planktonic offspring and adjust egg size accordingly. Instead, larvae may use information indicating risk in the plankton to adjust their size by cloning. Some cloned larvae were not much larger than conspecific eggs (Fig. 1B). Upon nearing metamorphosis, some clones were 2/3 the length (~1/3 the volume) of nonclones (Fig. 1C). Although size reduction may decrease vulnerability in the larval planktonic habitat, the persistence of small size implies that present risk for larvae is deferred to future risks for smaller juveniles on the sea floor.

Although larval cloning is known from all but one class of echinoderms (1), until now identified cues for cloning were limited to food and temperature favorable for growth and reproduction (3). Cloning in response to predators, like cloning in response to favorable conditions, can multiply each genet and increase fecundity, but advantages of predator-induced cloning do not depend on those features. If the small products of cloning are less vulnerable to predators, then cloning is adaptive whether all clonal products survive the process of division or only one.

Cloning by larvae in response to fish mucus suggests that even “restarting” development can provide a temporary escape from predation. We predict that cloning in response to predators will be found where safety lies in smaller size.

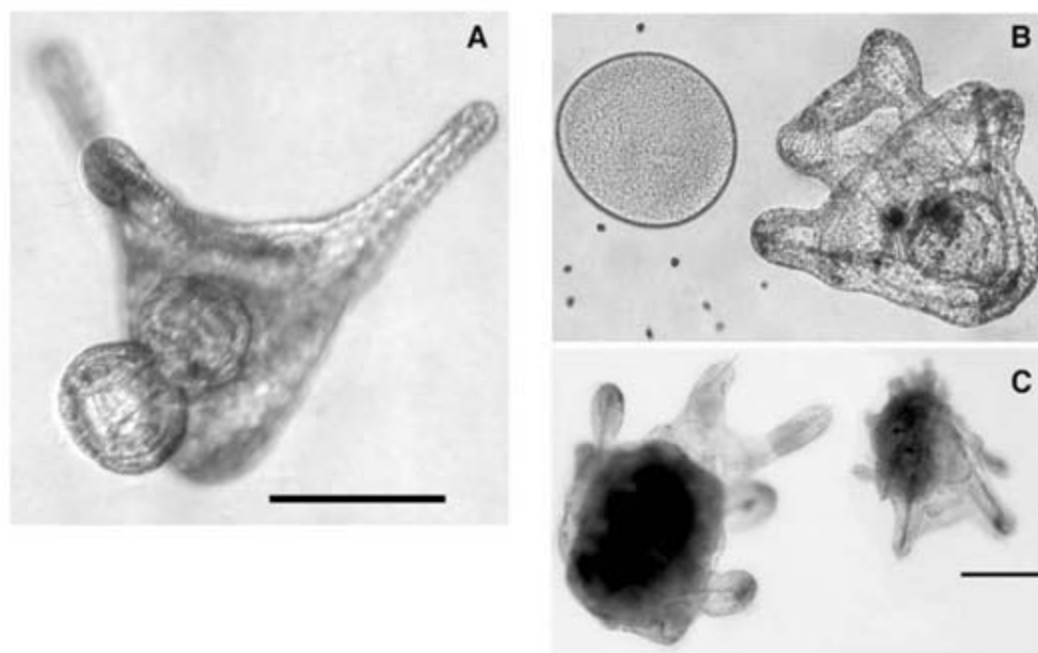


Fig. 1. Predator-induced cloning by early-stage sand dollar larvae (*Dendraster excentricus* plutei) and the effects of cloning on larval size. (A) A bud formed on the aboral surface of the primary larva that will subsequently be released as a gastrula-like individual. (B) A clone from asexual budding has developed to the six-arm stage but is only slightly larger than an egg (diameter of egg = 125 μ m). (C) A late-stage clone with juvenile rudiment (right) and a sand dollar larva of the same developmental stage that was not exposed to fish-mucus stimuli (left). Scale bars indicate 100 μ m in (A) and 200 μ m in (C).

References and Notes

1. A. A. Eaves, A. R. Palmer, *Nature* **425**, 146 (2003).
2. W. B. Jaekle, *Biol. Bull.* **186**, 62 (1994).
3. M. S. Vickery, J. B. McClintock, *Biol. Bull.* **199**, 298 (2000).
4. H. C. J. Godfray, *Parasitoids: Behavioral and Evolutionary Ecology* (Princeton Univ. Press, Princeton, NJ, 1994).
5. Materials and methods are available on Science Online.
6. M. K. Litvak, W. C. Leggett, *Mar. Ecol. Prog. Ser.* **81**, 13 (1992).
7. J. D. Allen, *Biol. Bull.* **214**, 42 (2008).
8. C. L. Suchman, B. K. Sullivan, *J. Plankton Res.* **22**, 2289 (2000).
9. We thank K. A. McDonald, F. X. Oyarzun, and M. F. Strathmann for discussion; L. Britt for collecting the fish predators used in this study; and K. Vaughn for assistance. Funding was provided by the Friday Harbor Laboratories and NSF (grant OCE0623102).

Supporting Online Material

www.sciencemag.org/cgi/content/full/319/5869/1503/DC1

Materials and Methods

Reference

22 October 2007; accepted 3 January 2008

10.1126/science.1151995

Friday Harbor Laboratories, University of Washington, 620 University Road, Friday Harbor, WA 98250, USA.

*To whom correspondence should be addressed. E-mail: dvaughn@u.washington.edu

Organic Molecules and Water in the Planet Formation Region of Young Circumstellar Disks

John S. Carr¹ and Joan R. Najita²

The chemical composition of protoplanetary disks is expected to hold clues to the physical and chemical processes that influence the formation of planetary systems. However, characterizing the gas composition in the planet formation region of disks has been a challenge to date. We report here that the protoplanetary disk within 3 astronomical units of AA Tauri possesses a rich molecular emission spectrum in the mid-infrared, indicating a high abundance of simple organic molecules (HCN, C₂H₂, and CO₂), water vapor, and OH. These results suggest that water is abundant throughout the inner disk and that the disk supports an active organic chemistry.

Disks of gas and dust orbit very young stars and represent the early stages of planetary system formation. Studies of these disks can establish the conditions under which planets form, complementing studies of solar system bodies in efforts to reconstruct the origin of our Solar System. In the same way that the chemical record preserved in primitive bodies (comets and meteorites) is probed for clues to the physical and chemical processes that operated in the early solar nebula (1, 2), the chemical composition of protoplanetary disks is expected to hold clues to the processes that are active in disks. Although significant advances have been made from the study of the dust properties of disks (3–5), much less progress has been made in the study of the gaseous component, particularly over the range of radii where planet formation is thought to occur (6–8). Measurements of disk molecular abundances can provide insights on issues such as the origin and evolution of water in the Solar System, the nature of organic chemistry in disks, and the ability of disks to synthesize pre-biotic species. Here we report that a protoplanetary disk around a young star displays a rich molecular emission spectrum in the mid-infrared that provides detailed thermal and chemical information on gas in the planet formation region of the disk.

We observed AA Tauri, a fairly typical classical T Tauri star (CTTS, which are approximately solar-mass stars with accretion disks and ages of less than a few 10⁶ years), with the Infrared Spectrograph (IRS) on the Spitzer Space Telescope (9, 10), covering wavelengths of 9.9 to 37.2 μm at a spectral resolution ($\lambda/\Delta\lambda$) of 600. The high signal-to-noise ratio (~200 to 300) produced by our observational and data reduction techniques (11) reveals a rich spectrum of molecular emission lines that is dominated by rotational trans-

sitions of H₂O (Fig. 1). Rotational transitions of OH were also observed between 20 and 31 μm. From 13 to 15 μm, ro-vibrational emission bands of the fundamental bending modes of C₂H₂, HCN, and CO₂ are prominent (Fig. 2).

This mid-infrared molecular emission almost certainly has its origin in a disk. AA Tauri lacks a surrounding molecular envelope or molecular outflow as potential alternate sites for the emitting gas. A disk interpretation is also consistent with high-spectral-resolution studies of CTTSs, which demonstrate a disk origin for their near-infrared molecular emission features. These include emission from the 5-μm fundamental ro-vibrational bands of CO, which is very common in CTTSs (12), and emission from the 2.3-μm CO overtone bands and hot bands of H₂O and OH (6, 13, 14), which is less frequently observed. All these near-infrared emission bands are thought to arise in a temperature inversion in the atmosphere of the inner (<1 astronomical unit (AU)) disk (6, 12–14).

To determine the properties of the emitting gas, we modeled the emission spectrum and analyzed the measured line fluxes using standard techniques (11). The emission from each molec-

ular species was modeled independently as thermally excited emission from a slab with a given temperature, column density, and emitting area. Although these quantities are expected to vary with radius in a real disk, we cannot constrain their radial variation without the benefit of velocity-resolved line profiles. Our strategy allows us to determine the flux-weighted mean gas temperature (T), column density (N), and emitting area ($\pi \times R^2$) for each molecule (Table 1). The combined synthetic spectrum reproduces well the details of the observed spectrum (Fig. 2). Molecular abundances with respect to CO (Table 1) were then determined by modeling the 5-μm fundamental CO emission spectrum of AA Tauri, obtained with the near-infrared echelle spectrograph NIRSPEC (15) at the Keck Observatory.

The equivalent radii for the emitting areas (Table 1) indicate that the observed gas resides at disk radii that correspond to the region of terrestrial planets in the Solar System; that is, within 2 to 3 AU of the star. The derived temperatures for the emission are also roughly consistent with predicted gas temperatures at radii ~1 AU for models that calculate the gas thermal structure in the temperature inversion region of disk atmospheres (16, 17). The significantly lower temperature for CO₂ suggests that its emission is concentrated at larger radii, which is also consistent with predictions of chemical models (18) that CO₂ is destroyed at temperatures much higher than 300 K. The higher mean temperature for CO can be understood by the fact that CO fundamental emission is insensitive to gas with $T < 400$ K, because of the higher energy of the upper levels of the transitions and the shorter wavelength of emission. Vertical gradients in the gas temperature and molecular abundances could also contribute to some of the differences; for example, H₂O may be abundant in a cooler part of the atmosphere as compared to CO (16).

The abundances of simple organics and H₂O for AA Tauri (Fig. 3) are about one order of magnitude higher than both observations and models (19) of hot molecular cores, the dense and warm

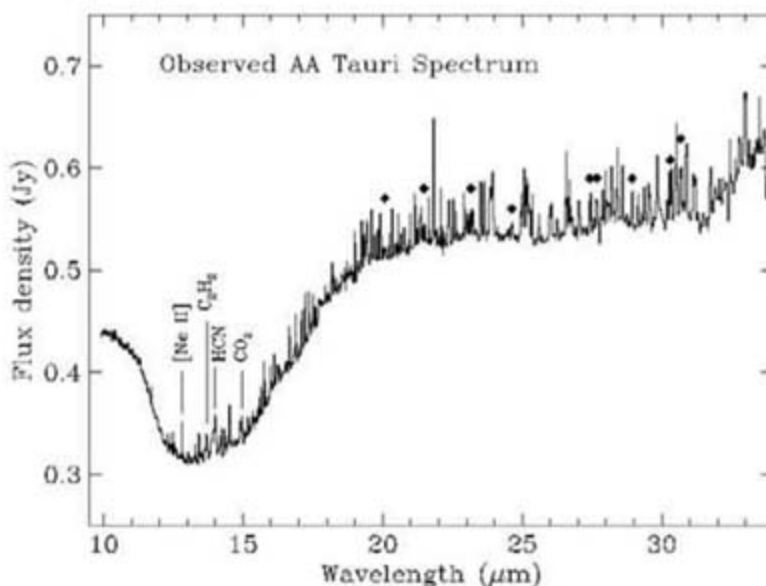


Fig. 1. Spitzer IRS spectrum of AA Tauri from 9.8 to 38.0 μm. The spectrum combines all orders of the short-high and long-high IRS modules (11). Rotational transitions of OH are marked with a diamond, and the Q branches of C₂H₂, HCN, and CO₂ are labeled along with atomic [Ne II]. The majority of emission features are due to rotational transitions of H₂O (unmarked features).

¹Remote Sensing Division, Naval Research Laboratory, Code 7210, Washington, DC 20375, USA. ²National Optical Astronomy Observatory, Tucson, AZ 85719, USA.

(several hundred degrees kelvin) regions of molecular gas that surround luminous protostars. The abundances of HCN and C₂H₂ in AA Tauri, relative to H₂O, are also substantially higher than measurements for cometary volatiles (fig. S1), whereas the abundance of CO₂ is consistent with or possibly less than its abundances in both comets and interstellar ices.

If molecular cores are representative of the composition of material that is incorporated into disks, then the higher abundances for AA Tauri

suggest that substantial molecular synthesis takes place within the disk. The observed region of the disk lies well inside the radius where ices would sublimate, based on the gas temperatures and inferred radii. This situation is similar to that of hot molecular cores, where a warm gas phase follows the sublimation of molecules from icy grain mantles. The subsequent gas-phase chemistry (19) may produce the complex organic molecules observed with millimeter-wave spectroscopy in both massive hot cores and the corresponding regions

surrounding low-mass protostars (20). Analogously, the indication of molecular synthesis in AA Tauri suggests that more complex organic molecules, including those of pre-biotic interest, might be produced within disks.

Similarly high abundances of C₂H₂, HCN, and CO₂ have also been observed (7) in absorption toward the low-mass protostar IRS 46, where the molecules potentially reside in a disk viewed at a favorable edge-on inclination or in an outflowing disk wind. Such gas absorption is rare toward low-mass protostars (7). In contrast, emission from a disk, such as the ubiquitous CO fundamental emission from CTTSs, does not require a special orientation to be observable. The fact that the molecular features in the mid-infrared spectrum of AA Tauri are detected in emission, combined with the fact that it is a typical CTTS, suggests the potential to study organic molecules and water in a large number of CTTS disks.

The high abundances in the AA Tauri disk are generally consistent with the predicted abundances at approximately 1-AU radii in a disk chemical model (21) for the inner 5 AU of protoplanetary disks (Fig. 3). However, the model results are vertically integrated abundances for the disk, whereas the molecular emission lines we observed are likely to probe only the upper disk atmosphere. The high abundances we derived are not expected in recent models that calculate the vertical thermal-chemical structure of the gas in disk atmospheres (16), which predict low abundances for molecules such as H₂O in the temperature inversion region of the atmosphere where emission lines form. If these models are correct, our measurements may be a sign of vertical mixing that carries molecular species from deep in the disk, where they are abundant, up to the surface where they are observed. Such turbulent mixing is predicted (22) by the magnetorotational instability (23), the process that is hypothesized to power disk accretion.

Our results for AA Tauri demonstrate the potential to measure the distribution of water vapor throughout the terrestrial planet region of disks, information that was previously restricted to hot H₂O inside of 0.3 AU (6, 13). The abundance of water vapor in the inner disk is important in setting the oxidation state of the gas, which influences the chemistry and mineralogy. Knowledge of the distribution of water (gas and ice) is also central to understanding the origin of water in inner solar system bodies and the formation of giant planets via the core accretion mechanism. The high abundance of H₂O for AA Tauri, if typical, would rule out simple models where outward diffusion rapidly depletes water vapor in the inner disk (24), leading to large ice density enhancements at larger radii and the rapid growth of giant planet cores. More detailed models (25) that include a variety of physical processes predict a more complicated evolution of water vapor in the inner disk, with both enhancements and depletions that vary with time and radius, indicating the importance of measuring the abundance and distribution of water with evolutionary age.

Fig. 2. Comparison of the observed spectrum of AA Tauri to the combined model spectrum (11). The observed continuum-subtracted spectrum from 13 to 16 μm (above) is offset from the model spectrum (below). The error bar indicates the average uncertainty in the observed spectrum. The model is a slab in local thermodynamic equilibrium, with an independent temperature, column density, and disk area derived for each molecular species (Table 1). All the unlabeled features are rotational transitions of H₂O.

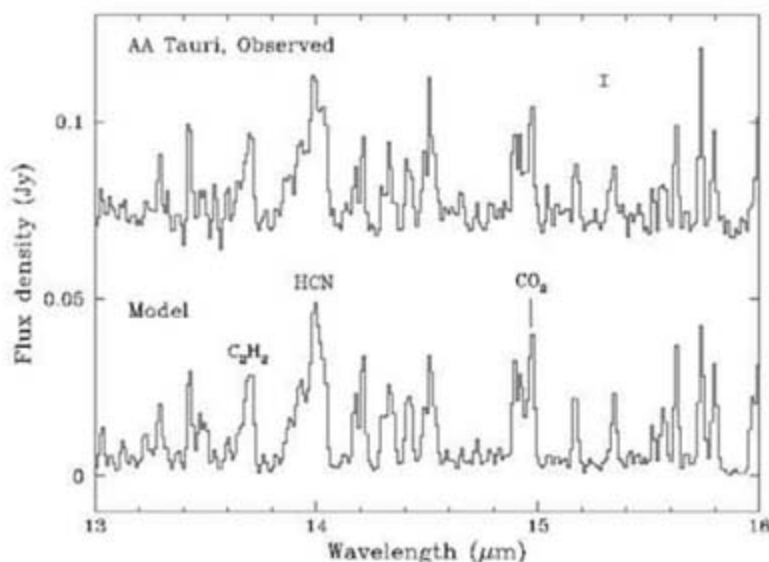


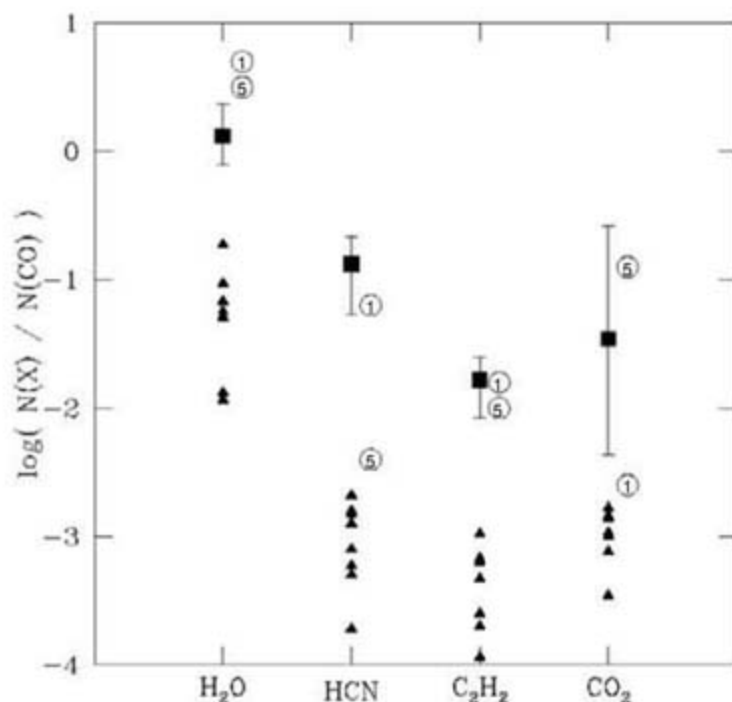
Table 1. Molecular gas parameters and abundances derived for AA Tauri.

Molecule	T (K)	N (10^{16} cm^{-2})	R^* (AU)	Abundance to CO
H ₂ O	575 ± 50	65 ± 24	2.1 ± 0.1	1.3
OH	525 ± 50	8.1 ± 5.2	2.2 ± 0.1	0.18
HCN	650 ± 100	6.5 ± 3.3	0.60 ± 0.05	0.13
C ₂ H ₂	650 ± 150	0.81 ± 0.32	0.60†	0.016
CO ₂	350 ± 100	0.2–13	1.2 ± 0.2	0.004–0.26
CO	900 ± 100	49 ± 16	0.7 ± 0.1	1.0

*The equivalent radius for the emitting area A ($R = [A/\pi]^{1/2}$).

†Area was set to that derived for HCN.

Fig. 3. Comparison of abundances relative to CO. The abundances for each molecule X relative to CO [$N(X)/N(\text{CO})$] derived for AA Tauri (squares) are compared to the abundances for hot molecular cores (triangles) and to disk chemical models (21) at radii of 1 and 5 AU (open circles with radius labeled). The abundances for hot cores are based on observations of absorption bands made with the Infrared Space Observatory (26–28). The organic molecules were studied using the same mid-infrared bands that were analyzed for AA Tauri.



References and Notes

- F. J. Ciesla, S. B. Chamley, in *Meteorites and the Early Solar System II*, D. S. Lauretta, H. Y. McSween Jr., Eds. (Univ. of Arizona Press, Tucson, AZ, 2006), pp. 209–230.
- D. Bockelée-Morvan, J. Crovisier, M. J. Mumma, H. A. Weaver, in *Comets II*, M. C. Festou, H. U. Keller, H. A. Weaver, Eds. (Univ. of Arizona Press, Tucson, AZ, 2004), pp. 391–423.
- E. Furlan *et al.*, *Astrophys. J. Suppl. Ser.* **165**, 568 (2006).
- J. E. Kessler-Silacci *et al.*, *Astrophys. J.* **639**, 275 (2006).
- F. J. Ciesla, *Science* **318**, 613 (2007).
- J. R. Najita, J. S. Carr, A. E. Glassgold, J. A. Valenti, in *Protostars and Planets V*, B. Reipurth, D. Jewitt, K. Keil, Eds. (Univ. of Arizona Press, Tucson, AZ, 2007), pp. 507–522.
- F. Lahuis *et al.*, *Astrophys. J.* **636**, L145 (2006).
- A. Dutrey, S. Guilloteau, P. Ho, in *Protostars and Planets V*, B. Reipurth, D. Jewitt, K. Keil, Eds. (Univ. of Arizona Press, Tucson, AZ, 2007), pp. 495–506.
- J. R. Houck *et al.*, *Astrophys. J. Suppl. Ser.* **154**, 18 (2004).
- M. W. Werner *et al.*, *Astrophys. J. Suppl. Ser.* **154**, 1 (2004).
- Materials and methods are available as supporting material on Science Online.
- J. Najita, J. S. Carr, R. D. Mathieu, *Astrophys. J.* **589**, 931 (2003).
- J. S. Carr, A. T. Tokunaga, J. Najita, *Astrophys. J.* **603**, 213 (2004).
- J. R. Najita, J. S. Carr, A. E. Glassgold, F. H. Shu, A. T. Tokunaga, *Astrophys. J.* **462**, 919 (1996).
- I. S. McLean *et al.*, *Proc. SPIE* **4008**, 1048 (2000).
- A. E. Glassgold, J. R. Najita, J. Igea, *Astrophys. J.* **615**, 972 (2004).
- H. Nomura, Y. Aikawa, M. Tsujimoto, Y. Nakagawa, T. J. Millar, *Astrophys. J.* **661**, 334 (2007).
- K. Willacy, H. H. Klahr, T. J. Millar, Th. Henning, *Astron. Astrophys.* **338**, 995 (1998).
- H. Nomura, T. J. Millar, *Astron. Astrophys.* **414**, 409 (2004).
- C. Ceccarelli, P. Caselli, E. Herbst, A. G. G. M. Tielens, E. Caux, in *Protostars and Planets V*, B. Reipurth, D. Jewitt, K. Keil, Eds. (Univ. of Arizona Press, Tucson, AZ, 2007), pp. 47–62.
- A. J. Markwick, M. Ilgner, T. J. Millar, Th. Henning, *Astron. Astrophys.* **385**, 632 (2002).
- N. J. Turner, K. Willacy, G. Bryden, H. W. Yorke, *Astrophys. J.* **639**, 1218 (2007).
- S. Balbus, J. Hawley, *Astrophys. J.* **376**, 214 (1991).
- D. J. Stevenson, J. I. Lunine, *Icarus* **75**, 146 (1988).
- F. J. Ciesla, J. N. Cuzzi, *Icarus* **181**, 178 (2006).
- F. Lahuis, E. F. van Dishoeck, *Astron. Astrophys.* **355**, 699 (2000).
- A. M. S. Boonman, E. F. van Dishoeck, F. Lahuis, S. D. Doty, *Astron. Astrophys.* **399**, 1063 (2003).
- A. M. S. Boonman, E. F. van Dishoeck, *Astron. Astrophys.* **403**, 1003 (2003).
- This work is based on observations made with the Spitzer Space Telescope, which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA. Support for this work was provided by NASA and by 6.1 base funding for basic research in infrared astronomy at the Naval Research Laboratory. Some of the data presented here were obtained at the W. M. Keck Observatory from telescope time allocated to NASA through the agency's scientific partnership with the California Institute of Technology and the University of California. The observatory was made possible by the generous financial support of the W. M. Keck Foundation.

Supporting Online Material

www.sciencemag.org/cgi/content/full/319/5869/1504/DC1

Materials and Methods

Figs. S1 and S2

References

5 December 2007; accepted 25 January 2008

10.1126/science.1153807

Superconductivity in Hydrogen Dominant Materials: Silane

M. I. Erements,^{1*} I. A. Trojan,^{1†} S. A. Medvedev,¹ J. S. Tse,² Y. Yao²

The metallization of hydrogen directly would require pressure in excess of 400 gigapascals (GPa), out of the reach of present experimental techniques. The dense group IVa hydrides attract considerable attention because hydrogen in these compounds is chemically precompressed and a metallic state is expected to be achievable at experimentally accessible pressures. We report the transformation of insulating molecular silane to a metal at 50 GPa, becoming superconducting at a transition temperature of $T_c = 17$ kelvin at 96 and 120 GPa. The metallic phase has a hexagonal close-packed structure with a high density of atomic hydrogen, creating a three-dimensional conducting network. These experimental findings support the idea of modeling metallic hydrogen with hydrogen-rich alloy.

The transformation of molecular hydrogen to a metal has been the subject of intense theoretical and experimental studies (1–6). However, the task of reaching the state of metallic hydrogen directly lies outside the possibilities of current techniques, requiring pressures in excess of ~400 GPa (7, 8). To circumvent this problem, it was proposed to compress hydrogen-rich compounds such as CH_4 and SiH_4 (9), in which the electron density on the hydrogen atoms is equivalent to pure hydrogen compressed to megabar pressures and therefore they are expected to metallize at much lower pressures. For the same reasons as for metallic hydrogen, very high Debye temperature and a strong electron-phonon interaction are expected, and it has been proposed that the metallic hydrides are candidates for high-

temperature superconductors. This suggestion has motivated considerable theoretical activity (10–15). First-principle calculations have predicted several probable structures for silane, SiH_4 , and different estimates for metallization pressures starting from as low as 20 to 90 GPa, with very high critical temperature for superconducting transition of $T_c \approx 40$ to 260 K (10–12). Metallization for GeH_4 (13) and superconductivity of SnH_4 with T_c close to 80 K at 120 GPa for SnH_4 (15) have also been predicted.

Experimental data on the hydrogen-rich compounds at high pressure are scarce. X-ray diffraction shows that at ambient pressure solid silane has a SnBr_4 -type structure and remains in molecular state at pressures at least up to 25 GPa (16). Reflection and transmission experiments indicate that at very high pressures of 288 GPa, methane displays semiconductor properties (17), although calculations show no metallization within the experimentally accessible pressure range (14).

We report here extensive studies of 10 samples of silane at pressures up to nearly 200 GPa using Raman scattering, electrical resistance mea-

surements, optical absorption, and x-ray powder diffraction. Starting from a few GPa, the Raman spectra show the characteristic bands around 2200 cm^{-1} and 900 cm^{-1} , corresponding to ν_3 , ν_1 , and ν_4 modes of internal vibrations of molecular SiH_4 , in accordance with known data at ambient pressure (18) (Fig. 1A). Discontinuities in the pressure dependence of Raman frequencies at 5 to 7 GPa and near 25 GPa indicate possible phase transitions (16). Between 50 and 65 GPa, the sample suddenly darkened and Raman spectra could no longer be detected. Electrical resistance measurements (Fig. 2A) showed that the sample resistance at room temperature dropped sharply, indicating the transformation to a metal (19). On cooling, a typical metallic behavior of the resistance was observed and eventually becoming superconducting (SC) at $T_c \approx 7$ K (Fig. 2B). Upon further compression, the sample became completely opaque at 76 GPa, and T_c increased, with pressure up to 17.5 K at 96 GPa and 17 K at 120 GPa (Fig. 2C). At higher pressures, T_c decreases to 8.8 K at 165 GPa and is then likely to increase again to 11.3 K at 192 GPa (Fig. 2C). The behavior of T_c between 90 GPa and 120 GPa is suggestive that higher values of critical temperature of superconductivity may be possible. However, uncontrollable change of pressure during sample loading (20) prohibited us from studying this regime in detail.

We clarified that the SC could not be affected by the Pt electrodes (fig. S1). Neither can the SC be attributed to pure silicon (20) as a result of possible decomposition of silane (11). We avoided decomposition by loading silane and performing measurements at low temperatures below 120 to 150 K. We found that the integrity of the sample is preserved when warmed to 300 K at pressures only above 100 GPa.

The observation of metallization and superconductivity in SiH_4 is in agreement with predictions (9). However, the pressure of metallization and the observed T_c values are different from

¹Max Planck Institute für Chemie, Postfach 3060, 55020 Mainz, Germany. ²Department of Physics and Engineering Physics, University of Saskatchewan, Saskatoon, S7N 5E2, Canada.

*To whom correspondence should be addressed. E-mail: erements@mpch-mainz.mpg.de

†On leave from A. V. Shubnikov Institute of Crystallography, Russian Academy of Sciences, 117333, Leninskii Avenue 59, Moscow, Russia.

predictions (10–12) for various structures. We therefore studied the structure of the sample. We indexed the diffraction pattern of the metallic phase at 113 GPa with a primitive hexagonal cell with lattice parameters $a = b = 2.67$ Å and $c = 4.49$ Å. The diffraction pattern is well described by a hexagonal close-packed (hcp) structure of Si atoms with two formula units per unit cell (Fig. 3A); The c/a ratio of 1.679 is close to the ideal hcp structure of $c/a = 1.633$. Although x-ray diffraction provides exact positions of Si atoms, the positions of hydrogen atoms cannot be determined from the diffraction pattern because of the very low atomic scattering factor. Fortunately, the close-packed arrangement of Si atoms proposes the unique set of positions for eight H atoms in the interstices of the unit cell. In the hcp structure, there are only four tetrahedral interstices and two octahedral ones, which are almost twice as large as the tetrahedral cages. Therefore,

the only way to place eight hydrogen atoms in the interstices of the unit cell is to have four hydrogen atoms occupy all tetrahedral interstices and the other four H atoms located in pairs in the large octahedral cages (Fig. 3C) (21). The resulting $P6_3$ structure is a compact packing of silicon and hydrogen atoms with covalent radii 1.11 Å and 0.37 Å, respectively. This structure can be considered as an ordered interstitial alloy of 20% of Si and 80% of hydrogen.

The resulting sublattice of hydrogen atoms is notable: In the ab -plane, they form regular hexagonal layers with an H-H distance of 1.54 Å, connected along the c -axis with an interlayer distance of 1.18 Å (Fig. 3C). Even the longest H-H distance within hexagons is shorter than the nearest H-Si distance (1.65 Å) (22). Thus, hydrogen atoms in the metallic $P6_3$ structure form a regular three-dimensional (3D) sublattice with considerable electronic hybridization due to short

H-H distances. The exact positions of hydrogen atoms could be determined with a neutron-scattering technique. However, the current technology limits the pressure under 40 GPa. Despite not knowing the exact positions of hydrogen atoms in the $P6_3$ phase, we can estimate an average density of hydrogen by subtracting the atomic volume of pure silicon from the sample volume measured at the same pressure (Fig. 4). At 60 GPa, it gives 4.5 Å³ for four hydrogen atoms, which corresponds to the volume per H₂ molecule of molecular hydrogen at 500 GPa (the pressures are obtained by extrapolation of equation of states (EOS) from (23)).

The metallic $P6_3$ phase is apparently stable in the 50 to 110 GPa range. However, if the sample is compressed to pressures ≥ 120 GPa, it transforms into a transparent phase, first appearing at the center of the sample, then spreading with time toward the edges. This transformation is slow,

Fig. 1. Raman scattering of silane at different pressures. (A) Representative Raman spectra in the low- and high-pressure regions taken at ambient temperature, with the exception of the 192 GPa spectrum, which was recorded at 20 K. Raman spectrum of the stressed diamonds changes from the sharp line at 1333 cm⁻¹ at the lowest pressures to the wide band with the high-frequency edge at 1690 cm⁻¹ at 192 GPa. (B) Pressure dependence of Raman frequencies. The lowest mode of the band at 2200 cm⁻¹ strongly softens. No Raman signal could be detected in the 50- to 110-GPa range because the sample became opaque. At higher pressures, a strong Raman signal appeared from a new phase. This phase has a large hysteresis, and a strong Raman signal persisted to low pressures upon releasing pressure (red points). (C) Comparison of experimental and calculated Raman spectra at 151 GPa. Asterisks denote features of the Raman band of stressed diamonds. The Raman modes below 400 cm⁻¹ are assigned to Si translation modes. The vibrations between

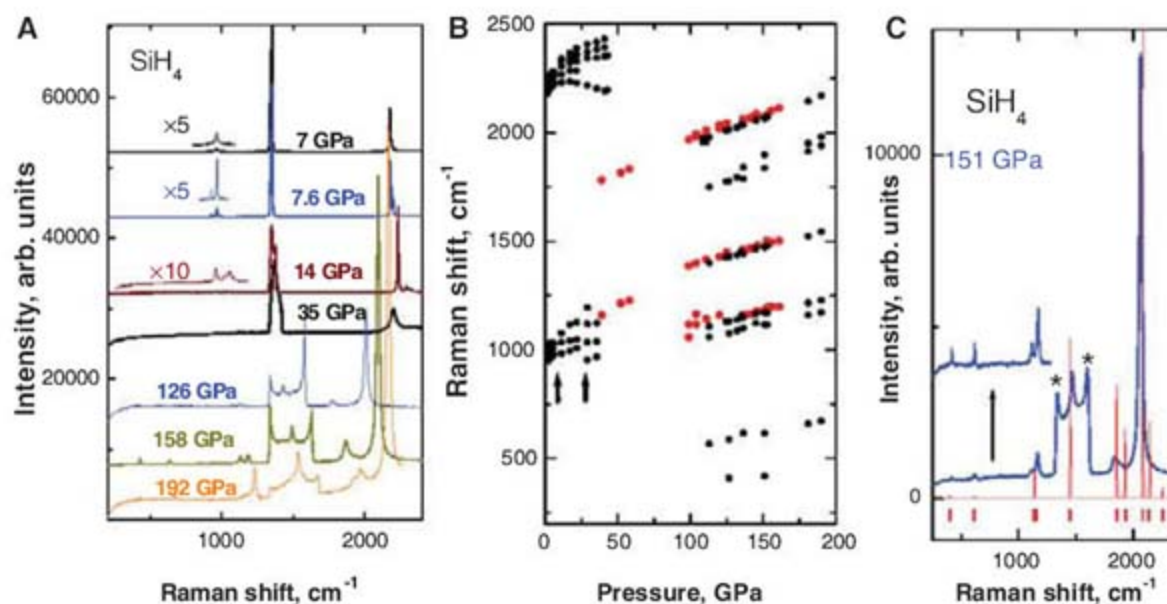
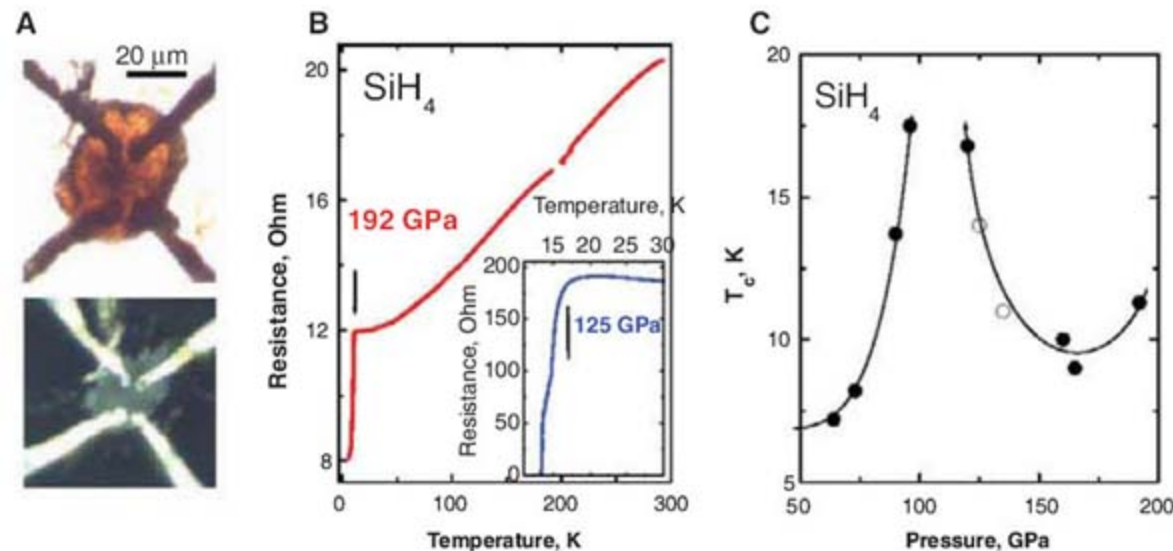


Fig. 2. Electrical resistance and superconductivity in silane. (A) Photographs of silane at 125 GPa taken in transmitted (above) and reflected (below) light after annealing at $T = 400^\circ\text{C}$. Silane is transparent at this pressure; nevertheless, it noticeably reflects light. (B) A representative superconducting step on the temperature dependence of resistance of silane at 192 GPa and 125 GPa (inset). Arrows indicate T_c . Electrical measurements were performed before the annealing, when an essential part of the sample remained in the metallic phase. (C) Pressure dependence of the critical transition temperature T_c of silane. The solid circles are data obtained when pressure was increasing, whereas the open circles show data obtained upon releasing pressure.



1000 and 1500 cm⁻¹ are attributed to H-Si-H bending modes. The high-frequency modes from 1800 to 2000 cm⁻¹ are localized Si-H stretching vibrations. The strongest Raman band at 2083 cm⁻¹ is assigned to an A_g Si-H stretch.

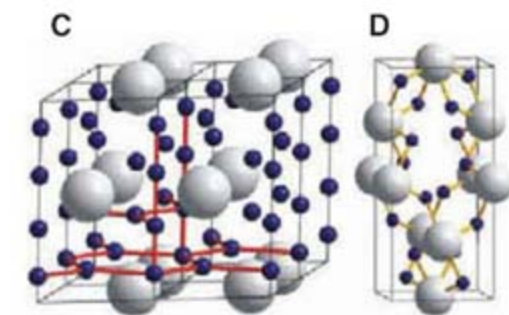
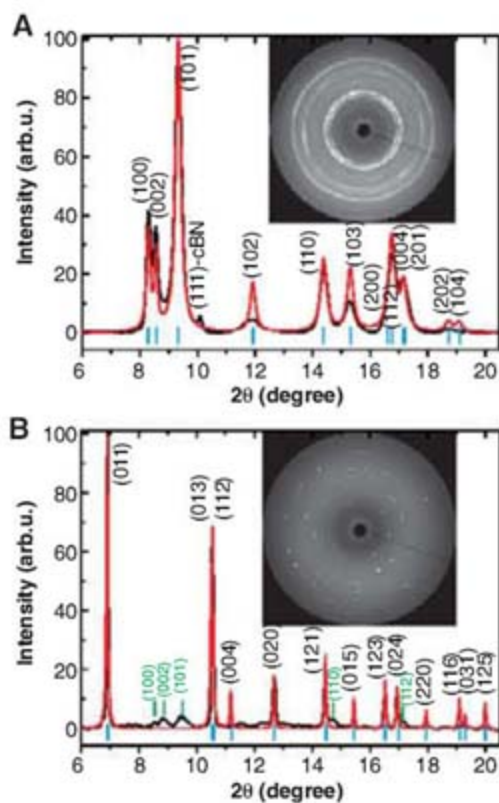


Fig. 3. X-ray diffraction patterns of metallic (A) and transparent (B) phases of SiH_4 at 113 GPa and 160 GPa, respectively. The black solid lines are experimental profiles; red solid lines are calculated [x-ray profiles were calculated using Diamond Crystal and Molecular Structure and Visualization software (Crystal Impact, Bonn, Germany)] for the $P6_3$ structure with lattice parameters $a = b = 2.671 \text{ \AA}$ and $c = 4.491 \text{ \AA}$, $V = 13.9 \text{ \AA}^3/\text{formula unit}$ (C) for the metallic phase and for the $I4_1/a$ structure with lattice parameters $a = b = 3.035 \text{ \AA}$ and $c = 6.817 \text{ \AA}$, $V = 15.7 \text{ \AA}^3/\text{formula unit}$ (D) for the transparent phase. The experimental integrated diffraction pattern in (B) was obtained by summarizing about 20 patterns measured from different areas of the sample. The tick marks below the profile show the

calculated reflection positions. Green ticks and labels in (B) indicate reflections from residual metallic phase coexisting at 160 GPa with $I4_1/a$ phase. Lattice parameters of metallic phase at this pressure are $a = b = 2.62 \text{ \AA}$ and $c = 4.39 \text{ \AA}$, $V = 13.1 \text{ \AA}^3/\text{formula unit}$. Gray spheres in (C) and (D) represent silicon atoms; the blue spheres are hydrogen atoms. The sublattice of hydrogen atoms in the inferred positions in the metallic $P6_3$ structure is shown by red lines. Both phases were measured at 113 GPa, 160 GPa, and 192 GPa.

developing over several days and involving only a part of the sample at room temperature. The transformation can be accelerated significantly by heating to $T > 400 \text{ K}$. After annealing for several hours, nearly the entire sample (except the edges of the sample and gasket) transformed to the transparent phase (Figs. 2A). Concurrent with the onset of transparency, a Raman signal appears and grows with time to a very pronounced spectrum (Fig. 1A). This spectrum appears at a pressure of $\approx 113 \text{ GPa}$ after annealing for several months at 450 K, while the sample remains completely opaque. Evidently, two distinct phases coexist to the highest experimental pressure of 192 GPa. In addition to the prominent Raman scattering, the transparent phase reveals a strong band at 2.45 eV in the optical absorption spectrum. When the pressure was released from 160 GPa, the transparent phase persisted at least down to 36 GPa, where the Raman signal became weaker (Fig. 1B).

The transparent phase produces sharp, spotty rings (Fig. 3B) in the x-ray diffraction patterns. We determined its structure to be a body-centered tetragonal, with lattice parameters $a = b = 3.03 \text{ \AA}$ and $c = 6.87 \text{ \AA}$ at 160 GPa. The experimental diffraction pattern can be described by a $I4_1/a$ structure, with Si atoms occupying the 4b positions (0, 0, 0.5) (Fig. 3B). The structure and its lattice parameters are in agreement with the $I4_1/a$ phase predicted to be thermodynamically stable and insulating in the pressure range 50 to

250 GPa (11). Following this theoretical work, we place the H atoms in the 16f positions (0.3676, 0.2166, 0.4328) (Fig. 3D). The Raman spectrum calculated for this structure with density-functional perturbation theory (24) is in agreement with the experiment (Fig. 1C). Therefore, the transparent $I4_1/a$ phase is undoubtedly stoichiometric silane. On the other hand, as described above, at low pressures molecular silane was stoichiometric, as the Raman spectra are typical for molecular SiH_4 (18). Therefore, under pressure the molecular phase transforms successively to the metallic $P6_3$ phase, then into the transparent $I4_1/a$ phase (Figs. 2A). Because there is no evidence for Si or H_2 (20), the intermediate metallic $P6_3$ phase should also be stoichiometric.

We now can understand the observed conductance and superconductivity behavior of silane at pressures above 120 GPa with the coexistence of the metallic phase and the transparent insulating $I4_1/a$ phase. The sharp increase of T_c with pressure of $\sim 100 \text{ GPa}$ (Fig. 2C) may be related to rapid changes in the Fermi surface that affect electron-phonon coupling substantially (e.g., due to nesting) at a narrow pressure range. Because a complete structural model for the $P6_3$ phase is not available, we will not speculate on the mechanism further.

The observed transformations in stoichiometric silane with pressure are unusual. Molecular SiH_4 at a pressure of $\approx 50 \text{ GPa}$ does not transform to the predicted thermodynamically stable $I4_1/a$

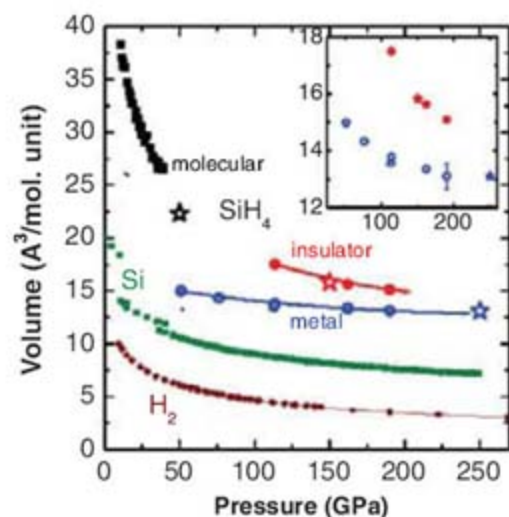


Fig. 4. EOS of silane, silicon (25), and molecular hydrogen (23). EOS of molecular silane at pressures below 25 GPa was taken from (16). At higher pressures, silane transforms to metal with $P6_3$ structure (blue points), then at $P > 110 \text{ GPa}$ an insulating $I4_1/a$ phase appears (red points), so that both phases coexist at higher pressures. Star points indicate theoretical calculations from (11) for molecular, insulating $I4_1/a$, and metallic $C2/c$ phase.

phase (11) but instead collapses to a significantly denser $P6_3$ phase (Fig. 4). According to theoretical calculations (11), a metallic $C2/c$ structure, which is structurally similar to $P6_3$ (21), is stable above 250 GPa. The appearance of the metallic phase at 50 GPa may be kinetically driven. It is likely that these two phases are indeed generically connected to each other because extrapolation of $V(P)$ for $P6_3$ phase to 250 GPa gives a volume close to the predicted value for theoretical $C2/c$ structure (Fig. 4).

Despite the high internal energy of the metallic phase, the Gibbs energy can be low. As the Gibbs free energy is the sum of three terms—the internal energy, the PV work, and the entropy—the overall stability of a higher internal energy metastable structure can be stabilized by having a much smaller volume and a large entropy. In this case, the $P6_3$ phase reported here indeed has a much smaller volume than the $I4_1/a$ phase. Moreover, the possibility of disordering in the H positions may also help to increase the entropy. The $P6_3$ and $I4_1/a$ phases can be separated with a significant kinetic barrier because of a large difference in volumes. This difference decreases with pressure (Fig. 4), and therefore transformation to $I4_1/a$ phase becomes possible. Interestingly, at the highest pressures of $\sim 200 \text{ GPa}$, this transformation slows down, apparently because volumes (Fig. 4) and energies of the two phases approach each other.

References and Notes

1. E. Wigner, H. B. Huntington, *J. Chem. Phys.* **3**, 764 (1935).
2. N. W. Ashcroft, *Phys. Rev. Lett.* **21**, 1748 (1968).
3. E. G. Brovman, Y. Kagan, A. Kholas, *Sov. Phys. JETP* **35**, 783 (1972).
4. H. K. Mao, R. J. Hemley, *Rev. Mod. Phys.* **66**, 671 (1994).
5. W. J. Nellis, *Rep. Prog. Phys.* **69**, 1479 (2006).

6. P. Loubeyre, F. Occelli, R. LeToullec, *Nature* **416**, 613 (2002).
7. K. A. Johnson, N. W. Ashcroft, *Nature* **403**, 632 (2000).
8. M. Stadelé, R. M. Martin, *Phys. Rev. Lett.* **84**, 6070 (2000).
9. N. W. Ashcroft, *Phys. Rev. Lett.* **92**, 187002 (2004).
10. J. Feng *et al.*, *Phys. Rev. Lett.* **96**, 017006 (2006).
11. C. J. Pickard, R. J. Needs, *Phys. Rev. Lett.* **97**, 045504 (2006).
12. Y. Yao, J. S. Tse, Y. Ma, K. Tanaka, *Eur. Phys. Lett.* **78**, 37003 (2007).
13. M. Martínez-Canales, A. Bergara, J. Feng, W. Grochala, *J. Phys. Chem. Solids* **67**, 2095 (2006).
14. M. Martínez-Canales, A. Bergara, *High Pressure Res.* **26**, 369 (2006).
15. J. S. Tse, Y. Yao, K. Tanaka, *Phys. Rev. Lett.* **98**, 117004 (2007).
16. O. Degtyareva *et al.*, *Phys. Rev. B* **76**, 064123 (2007).
17. L. Sun, A. L. Ruoff, C. S. Zha, G. Stupian, *J. Phys. Chem. Solids* **67**, 2603 (2006).
18. R. P. Fournier, R. Savoie, N. D. The, R. Belzile, A. Cabana, *Can. J. Chem.* **50**, 35 (1972).
19. It is accompanied by a large drop of volume, which in turn promotes a uniform pressure over the sample, as evidenced from the sharpening of the high-frequency edge of a Raman spectrum from a diamond anvil (Fig. 1).
20. Materials and methods are available as supporting material on Science Online.
21. Although for symmetry reasons, H atoms in the tetrahedral interstices occupy centers of the cavities, there is some freedom in the choice of z coordinates of H atoms in the octahedral cavities. We have chosen the z coordinates for all positions in such a way that all H atoms are located at equal distances from surrounding Si atoms (Fig. 3C). This arrangement of H atoms corresponds to $P6_3/mmc$ space group. However, because of the mentioned uncertainty of the coordinates of H atoms, further on we refer to this phase as the lowest consistent with hcp structure $P6_3$ symmetry, with Si atoms in the 2b positions (1/3, 2/3, 0) and H atoms in the positions 2a (0, 0, 0.131), 2a (0, 0, 0.369), 2b (2/3, 1/3, 0.131), and 2b (1/3, 2/3, 0.369) (because H atoms occupy all interstices of hcp structure, they are forced to conform to at least this symmetry).
22. In some aspects, $P6_3$ phase is close to the metallic C2/c structure predicted in (11). Because of the specific metric of this structure ($\alpha \approx c$, $\beta = 115.2^\circ$), Si atoms form slightly distorted close-packed layers in ac planes stacked along the b axis. The H atoms are located in three nonequivalent sites in which an H atom forms bonds to two, three, or four Si atoms. The last two positions are similar to octahedral and tetrahedral positions of H atoms in $P6_3$ structure, respectively. Besides, hydrogen atoms in C2/c structure form a 3D network with H-H distances 1.24 to 1.57 Å, which is very close to H-H distances in $P6_3$ structure.
23. P. Loubeyre *et al.*, *Nature* **383**, 702 (1996).
24. S. Baroni, S. de Gironcoli, A. Dal Corso, P. Giannozzi, *Rev. Mod. Phys.* **73**, 515 (2001).
25. S. J. Duclos, Y. K. Vohra, A. L. Ruoff, *Phys. Rev. B* **41**, 12021 (1990).
26. We are grateful to A. Gavriluk and V. Prakapenka for help in x-ray diffraction measurements at station Sector 13 Insertion Device Beamline (13ID-D) of the GeoSoilEnviro group of the Consortium for Advanced Radiation Sources (GSECARS) at the Advanced Photon Source under grant GUP7569 and GUP8324. We thank R. Boehler, Y. A. Freiman, I. Gondharenko, and V. Hilgner for fruitful discussions. Deutsche Forschungsgemeinschaft (DFG) grant ER 539/1-1 provided partial financial support.

Supporting Online Material

www.sciencemag.org/cgi/content/full/319/5869/1506/DC1
Materials and Methods

Fig. S1
References

26 November 2007; accepted 24 January 2008
10.1126/science.1153282

Energy Gaps and Kohn Anomalies in Elemental Superconductors

P. Aynajian,¹ T. Keller,^{1,2} L. Boeri,¹ S. M. Shapiro,³ K. Habicht,⁴ B. Keimer^{1*}

The momentum and temperature dependence of the lifetimes of acoustic phonons in the elemental superconductors lead and niobium were determined by resonant spin-echo spectroscopy with neutrons. In both elements, the superconducting energy gap extracted from these measurements was found to converge with sharp anomalies originating from Fermi-surface nesting (Kohn anomalies) at low temperatures. The results indicate electron many-body correlations beyond the standard theoretical framework for conventional superconductivity. A possible mechanism is the interplay between superconductivity and spin- or charge-density-wave fluctuations, which may induce dynamical nesting of the Fermi surface.

Over the past half-century, a comprehensive framework based on the Bardeen-Cooper-Schrieffer formulation (1) has been developed for the interpretation of experimental data on superconductors. Although this framework has been challenged by the discovery of high-temperature superconductivity, it provides a remarkably successful description of the physical properties of conventional low-temperature superconductors (2). Even today, however, the prediction of two of the most important quantities characterizing a superconductor, the transition temperature and the energy gap at the Fermi level, from first principles presents a formidable challenge to theory because they depend exponentially on material-specific parameters such as the

phononic and electronic densities of states and the electron-phonon coupling (3). We present neutron scattering data on the lifetimes of acoustic phonons in Pb and Nb (the two elements with the highest superconducting transition temperatures, $T_c = 7.2$ and 9.3 K, respectively) that shed light on the energy gap in conventional superconductors.

The energy gap can be directly determined in phonon lifetime measurements, because electron-phonon scattering is suppressed (and the phonon lifetimes are thus enhanced) for energies below the gap. Our data indicate a surprising relation between the superconducting gap and the geometry of the Fermi surface, which also leaves an imprint on the phonon lifetimes (4): For phonon wave vectors connecting nearly parallel segments of the Fermi surface, the electron-phonon scattering probability is enhanced, and lifetime extrema (termed Kohn anomalies) are generally expected. We have recorded hitherto unknown Kohn anomalies in both Pb and Nb and found that the low-temperature energy gap coincides with such an anomaly in both materials. This phenomenon has not been anticipated by the

standard theoretical framework for conventional superconductors.

Both Kohn anomalies (5–8) and superconductivity-induced phonon renormalization (9) have been observed by inelastic neutron scattering. However, because the requisite energy resolution is difficult to obtain, these investigations have been limited to a few selected materials, and both effects have thus far not been studied accurately in the same material. The systematic investigation reported here was made possible by recent advances in resonant spin-echo spectroscopy with neutrons (10–12), which have enabled the determination of the lifetimes of dispersive excitations with μeV energy resolution over the entire Brillouin zone. In brief, the spin echo is generated on a triple-axis spectrometer by using radio-frequency magnetic fields to manipulate the spin polarization of neutrons scattered from a crystal before and after the scattering event. The excitation lifetime is then extracted from the spin-echo decay profile.

The measurements were taken on high-purity Pb and Nb single crystals. The resulting spin-echo decay profiles for selected transverse acoustic phonons in Pb and Nb (Fig. 1) are well described by exponentials, corresponding to Lorentzian phonon spectral functions; deviations from Lorentzian line shapes were not found within the experimental error. The spin-echo decay rate (proportional to the phonon linewidth and inversely proportional to its lifetime) decreases upon lowering the temperature, reflecting the loss of the electron-phonon decay channel in the superconducting state. The nonzero decay rate at the lowest temperatures is due to instrumental limitations, which can be quantitatively determined on the basis of the phonon dispersion relations and the mosaic spreads of the single-crystal samples (13). The intrinsic Lorentzian phonon linewidths, Γ , are extracted by fitting the decay profiles

¹Max Planck Institute for Solid State Research, Heisenbergstrasse 1, D-70569 Stuttgart, Germany. ²ZWE FRM-II, Technical University of Munich, Lichtenbergstrasse 1, D-85748 Garching, Germany. ³Brookhaven National Laboratory, Upton, NY 11973–5000, USA. ⁴Hahn Meitner Institute, Glienicke Strasse 100, D-14109 Berlin, Germany.

*To whom correspondence should be addressed. E-mail: b.keimer@fkf.mpg.de

to exponentials (lines in Fig. 1) and correcting for this instrumental contribution.

The intrinsic wave vector (q)-dependent linewidths of the lowest-energy, transverse acoustic phonon mode T_1 of Pb along $q = (\xi, \xi, 0)$ is shown in Fig. 2A. At all temperatures, sharp anomalies in the phonon linewidths are seen at $\xi \sim 0.25, 0.35$, and 0.50 reciprocal lattice units (r.l.u.). Although the phonon spectrum of Pb has been studied extensively by conventional neutron spectroscopy, these particular features have not been recognized because of insufficient energy resolution. The same features also appear in the phonon dispersion relation (Fig. 2B): Max-

ima in the phonon linewidth coincide with characteristic S-shaped deviations from the q linear dispersion, as stipulated by the Kramers-Kronig relation that holds for all excitations in solids. Artefacts associated with the new measurement method would generally not be Kramers-Kronig consistent and can thus be ruled out. This implies that phonons with the anomalous wave vectors shown in Fig. 2A are intrinsically unstable toward decay into other elementary excitations. In principle, the decay products can be either other phonons (generated, for instance, by anharmonic terms in the lattice potential) or electron-hole pairs (originating from Kohn anomalies). The features at

$\xi \sim 0.35$ and 0.5 can be associated with Kohn anomalies because these wave vectors are known as nesting vectors of the Fermi surface (14). Indeed, Kohn anomalies have been observed at these wave vectors in the longitudinal phonon branch of Pb (5, 6, 13). The origin of the feature at $\xi \sim 0.25$ is more subtle because this wave vector does not match any known spanning vector of the Fermi surface. A possible origin is a three-phonon decay process previously observed in the spectrum of phonons in liquid helium, which are unstable because their phase velocity exceeds the velocity of sound (15, 16). Indeed, accurate measurements of the phonon dispersions in Pb (Fig. 2B) show that the phonon phase velocity exceeds the sound velocity around $\xi \sim 0.25$, presumably as a consequence of the dispersion anomaly at $\xi \sim 0.35$. This process has thus far not been observed in solids and deserves further investigation. Anharmonic terms in the lattice potential may also contribute to the anomaly.

We focused on the influence of superconductivity on the phonon linewidths below $T_c = 7.2$ K. As the superconductor is cooled below T_c , the electron-hole decay channel is closed (and Γ is reduced) below the energy gap $2\Delta(T)$. This effect is observed at low wave vectors ξ in Fig. 2A. In particular, Γ approaches 0 for $T \ll T_c$ around $\xi = 0.32$ [corresponding to a phonon energy of 2.47 meV, below the low-temperature limit of $2\Delta \sim 2.7$ meV known from tunneling measurements (17)]. For lower energies around $\xi \sim 0.25$, however, Γ remains non-zero even at the lowest temperatures, supporting the notion that the linewidth anomaly at this wave vector originates from the three-phonon down-conversion process discussed above and/or lattice anharmonicity and not from electron-hole pair production. We removed the contribution of this process for clarity and show only the phonon linewidth, Γ_{e-p} , directly attributable to the electron-phonon interaction (Fig. 3). As expected, Γ_{e-p} exhibits a maximum because of the pileup of

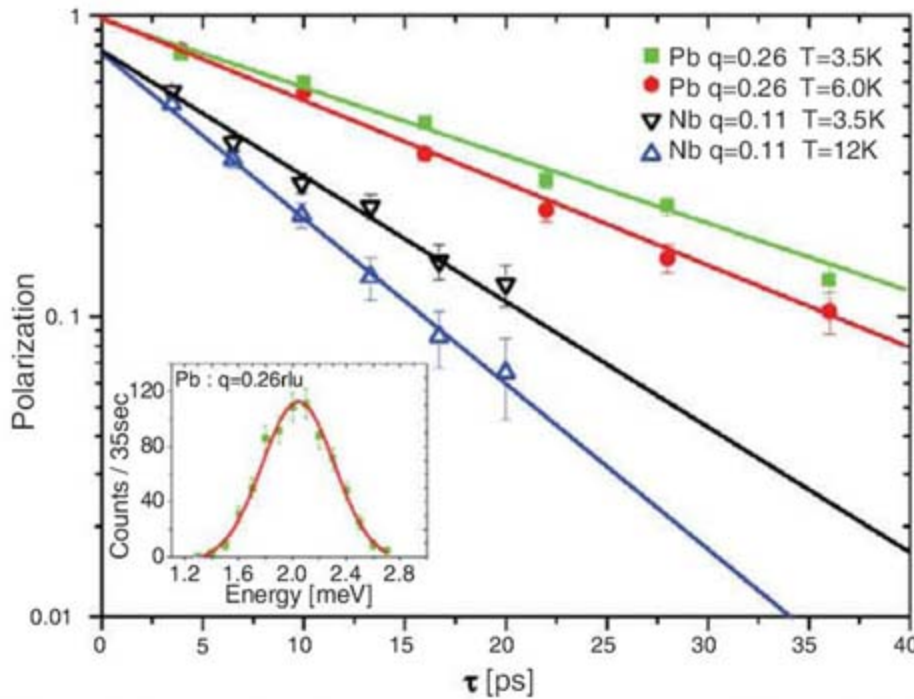


Fig. 1. Spin-echo decay profiles of transverse acoustic phonons at $q = (0.26, 0.26, 0)$, phonon energy $E = 2.32$ meV in Pb (top two curves) and $q = (0.11, 0, 0)$, $E = 2.06$ meV in Nb (bottom two curves) at selected temperatures. The spin polarization of the beam at the detector is plotted versus the spin-echo time τ (10, 11). The lines are the results of fits of exponentials (corresponding to Lorentzian spectral functions) to the data. (Inset) A conventional triple-axis scan through the phonon in Pb. Error bars indicate the statistical errors.

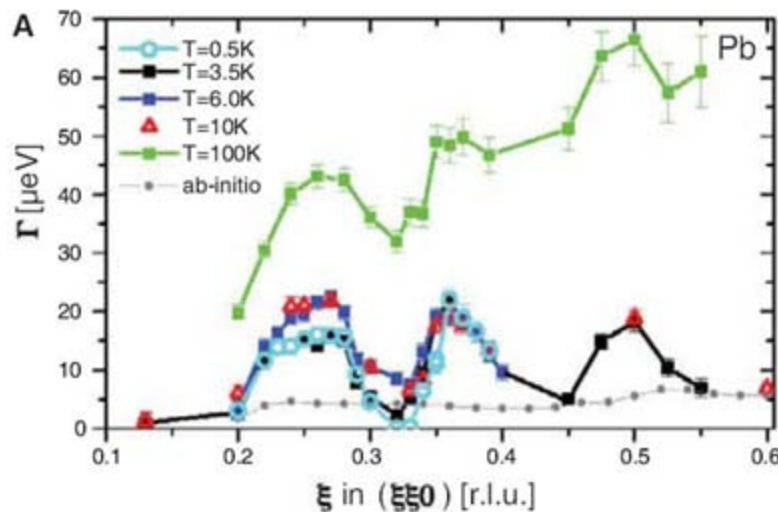
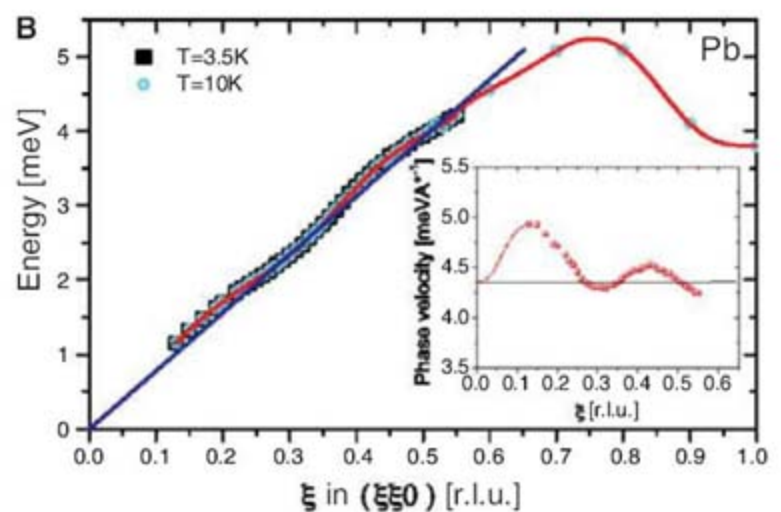


Fig. 2. (A) Linewidths of transverse acoustic phonons along $q = (\xi, \xi, 0)$ in Pb at selected temperatures. The data were obtained by correcting the measured spin-echo decay rates for instrumental effects (13). The gray symbols are the results of ab initio lattice-dynamical calculations, as described in the text. Error



bars indicate the statistical errors. **(B)** Dispersion relation of the same phonon extracted from triple-axis data. (Inset) The phonon phase velocity (E/q) computed from the data. The blue line in (B) and the black line in the (B) inset represent the experimentally determined sound velocity (29).

electronic density of states above $2\Delta(T)$, which moves to progressively higher energies upon cooling and closely tracks the energy gap determined in prior tunneling measurements (17) (Fig. 3 inset). Surprisingly, however, the superconductivity-induced maximum of Γ_{e-p} merges with the Kohn anomaly as $T \rightarrow 0$. At $T = 0.5$ K, both anomalies are indistinguishable within the measurement error.

In order to explore whether the coincidence of $2\Delta(T \rightarrow 0)$ and the Kohn anomaly in Pb is accidental, we performed similar experiments on phonons in Nb, an elemental superconductor with a different Fermi surface and phonon spectrum. Figure 4A shows the momentum-dependent linewidths of the transverse acoustic phonon branch along $q = (\xi, 0, 0)$ in Nb at temperatures above and below $T_c = 9.3$ K. The data above T_c are in fair agreement with prior work (9), but they reveal

several sharp features that have not been identified before. In part on the basis of ab initio lattice dynamical calculations, they can be identified as Kohn anomalies (see below). The existence of a Kohn anomaly at $\xi \sim 0.17$ persisting up to room temperature has been suggested on the basis of prior experimental work (7, 18). As described above for Pb, the linewidths are reduced below and enhanced above the gap for quasi-particle-pair production, $2\Delta(T)$, in the superconducting state, and the low-temperature electron-phonon linewidth shows the expected dependence on wave vector (or energy). Similar to the observation in Pb, the $2\Delta(T \rightarrow 0)$ extracted from the low-temperature Γ_{e-p} of Nb again coincides with the lowest-energy Kohn anomaly within the experimental error (Fig. 4B).

To help interpret these observations, we calculated the phonon dispersions and linewidths

in the framework of ab-initio density functional perturbation theory in the local-density approximation (LDA) (13) on a very fine mesh of q points in reciprocal space. The phonon frequencies were obtained by diagonalization of the dynamical matrices and the electron-phonon linewidths by Allen's formula (19). The results are in reasonable overall agreement with the experimental data (Figs. 3 and 4). In particular, both the phonon frequencies and the linewidths associated with Kohn anomalies in the high-energy transverse acoustic phonons of Nb (Fig. 4A) and in the longitudinal phonon of Pb (13) are well described, indicating that the resolution of the calculations is sufficient to reproduce subtle structures in q space.

The lowest-lying Kohn anomalies in the transverse-acoustic phonon branches of both Pb and Nb are, however, not reproduced by the calculations (Figs. 3A and 4, A and B). These anomalies therefore originate in factors not included in the calculations, such as the relativistic spin-orbit coupling, phonon nonadiabaticity (20), or many-body correlations beyond Allen's formula (21, 22) or the LDA. Because the Kohn anomalies in Pb and Nb are of comparable strength, the spin-orbit coupling (which is much stronger in Pb than in Nb) cannot be responsible. Because of the large Fermi energies of both materials, nonadiabatic electron-phonon coupling effects should also be extremely weak.

This leaves electron correlation effects beyond the LDA as the most likely mechanism responsible for the low-energy Kohn anomalies. It seems reasonable to assume that the same correlations are also responsible for the observed coincidence of $2\Delta(T \rightarrow 0)$ with the same anomalies. Because the anomalies persist to temperatures above 100 K, superconducting fluctuations are unlikely to be directly responsible. We note, however, that the formation of spin or charge density waves driven by electron correlations has been predicted for Pb and other elemental metals (23). Although extensive searches for static

Fig. 3. Contribution of the electron-phonon interaction to the linewidth of the transverse acoustic phonon along $q = (\xi, 0, 0)$ in Pb. The corresponding E is provided by the scale at the top; note the non-linear E -versus- q relationship (Fig. 2B). The lines are results of least-squares fits to the Bardeen-Cooper-Schrieffer (BCS) excitation spectrum function (30). Error bars indicate the statistical errors. (Inset) The temperature dependence of the superconducting energy gap (squares) and Kohn anomaly (triangles) extracted from the fits. The line in the inset shows the BCS expression for the superconducting gap (2), which was experimentally confirmed by tunneling spectroscopy (17). The error bars account for systematic errors associated with extracting the peak positions. At $T = 0.5$ K, for instance, fits to two peaks with energies at the end points of the error bar are in acceptable agreement with the data.

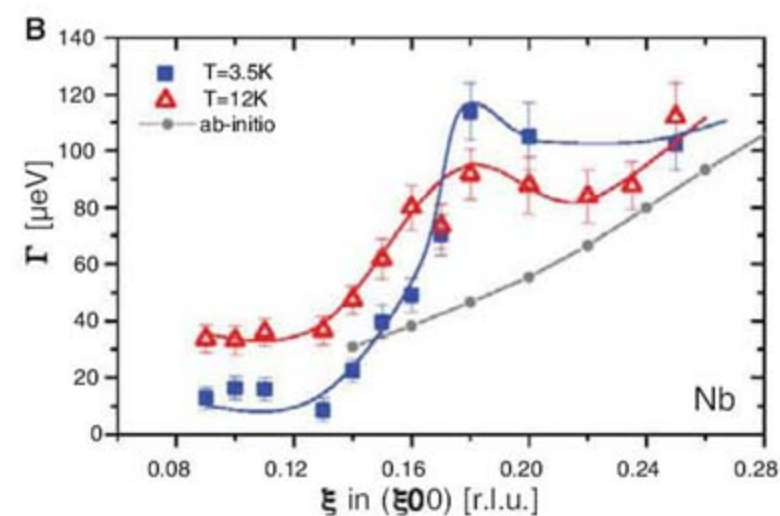
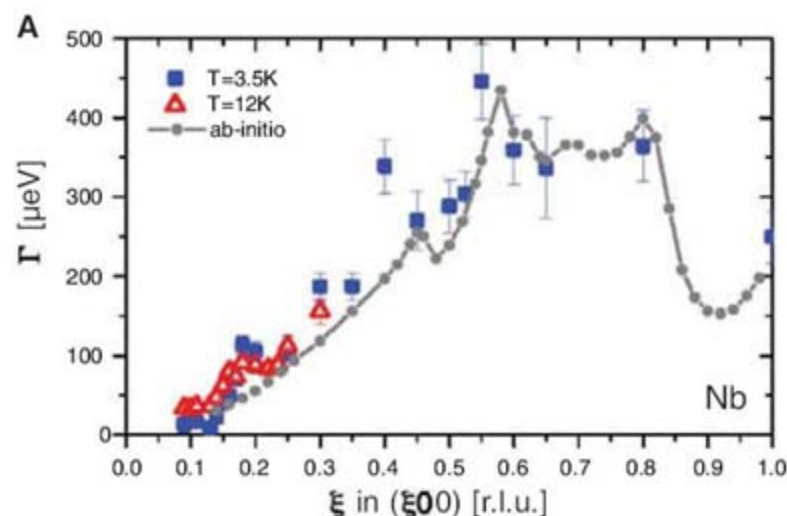
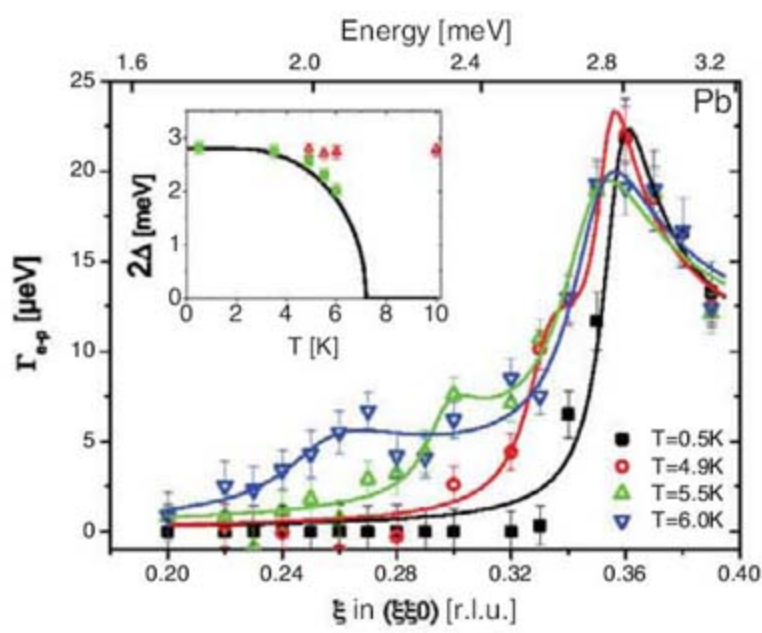


Fig. 4. (A) Linewidths of transverse acoustic phonons along $q = (\xi, 0, 0)$ in Nb at two different temperatures. The gray symbols are the results of lattice-dynamical calculations, as described in the text. (B) Blowup

of the low- q segment of (A). The corresponding E is provided by the scale at the top. The lines are guides to the eye. Error bars indicate the statistical errors.

to exponentials (lines in Fig. 1) and correcting for this instrumental contribution.

The intrinsic wave vector (q)-dependent linewidths of the lowest-energy, transverse acoustic phonon mode T_1 of Pb along $q = (\xi, \xi, 0)$ is shown in Fig. 2A. At all temperatures, sharp anomalies in the phonon linewidths are seen at $\xi \sim 0.25, 0.35$, and 0.50 reciprocal lattice units (r.l.u.). Although the phonon spectrum of Pb has been studied extensively by conventional neutron spectroscopy, these particular features have not been recognized because of insufficient energy resolution. The same features also appear in the phonon dispersion relation (Fig. 2B): Max-

ima in the phonon linewidth coincide with characteristic S-shaped deviations from the q linear dispersion, as stipulated by the Kramers-Kronig relation that holds for all excitations in solids. Artefacts associated with the new measurement method would generally not be Kramers-Kronig consistent and can thus be ruled out. This implies that phonons with the anomalous wave vectors shown in Fig. 2A are intrinsically unstable toward decay into other elementary excitations. In principle, the decay products can be either other phonons (generated, for instance, by anharmonic terms in the lattice potential) or electron-hole pairs (originating from Kohn anomalies). The features at

$\xi \sim 0.35$ and 0.5 can be associated with Kohn anomalies because these wave vectors are known as nesting vectors of the Fermi surface (14). Indeed, Kohn anomalies have been observed at these wave vectors in the longitudinal phonon branch of Pb (5, 6, 13). The origin of the feature at $\xi \sim 0.25$ is more subtle because this wave vector does not match any known spanning vector of the Fermi surface. A possible origin is a three-phonon decay process previously observed in the spectrum of phonons in liquid helium, which are unstable because their phase velocity exceeds the velocity of sound (15, 16). Indeed, accurate measurements of the phonon dispersions in Pb (Fig. 2B) show that the phonon phase velocity exceeds the sound velocity around $\xi \sim 0.25$, presumably as a consequence of the dispersion anomaly at $\xi \sim 0.35$. This process has thus far not been observed in solids and deserves further investigation. Anharmonic terms in the lattice potential may also contribute to the anomaly.

We focused on the influence of superconductivity on the phonon linewidths below $T_c = 7.2$ K. As the superconductor is cooled below T_c , the electron-hole decay channel is closed (and Γ is reduced) below the energy gap $2\Delta(T)$. This effect is observed at low wave vectors ξ in Fig. 2A. In particular, Γ approaches 0 for $T \ll T_c$ around $\xi = 0.32$ [corresponding to a phonon energy of 2.47 meV, below the low-temperature limit of $2\Delta \sim 2.7$ meV known from tunneling measurements (17)]. For lower energies around $\xi \sim 0.25$, however, Γ remains non-zero even at the lowest temperatures, supporting the notion that the linewidth anomaly at this wave vector originates from the three-phonon down-conversion process discussed above and/or lattice anharmonicity and not from electron-hole pair production. We removed the contribution of this process for clarity and show only the phonon linewidth, Γ_{e-p} , directly attributable to the electron-phonon interaction (Fig. 3). As expected, Γ_{e-p} exhibits a maximum because of the pileup of

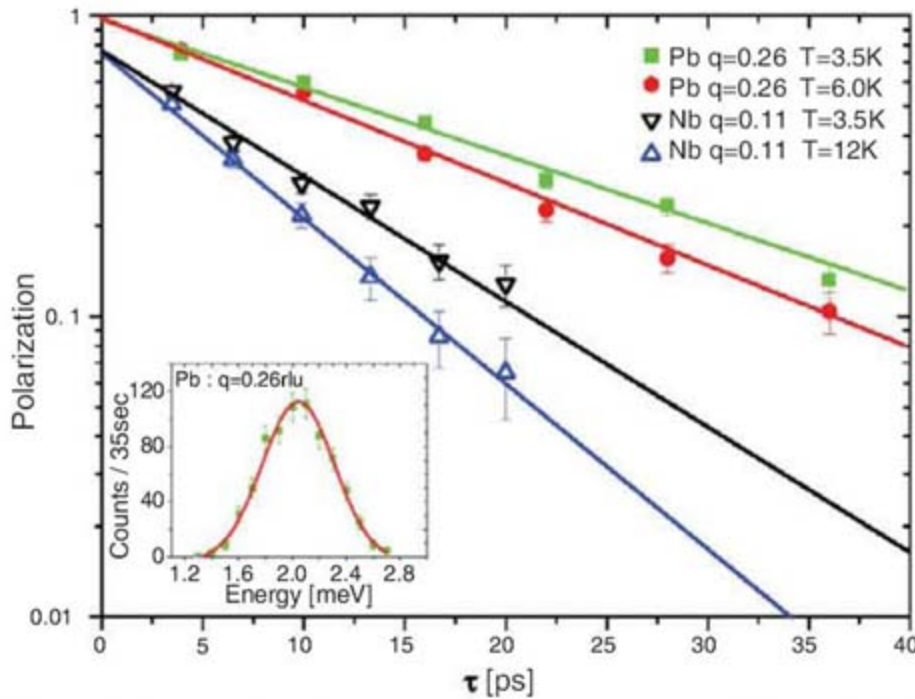


Fig. 1. Spin-echo decay profiles of transverse acoustic phonons at $q = (0.26, 0.26, 0)$, phonon energy $E = 2.32$ meV in Pb (top two curves) and $q = (0.11, 0, 0)$, $E = 2.06$ meV in Nb (bottom two curves) at selected temperatures. The spin polarization of the beam at the detector is plotted versus the spin-echo time τ (10, 11). The lines are the results of fits of exponentials (corresponding to Lorentzian spectral functions) to the data. (Inset) A conventional triple-axis scan through the phonon in Pb. Error bars indicate the statistical errors.

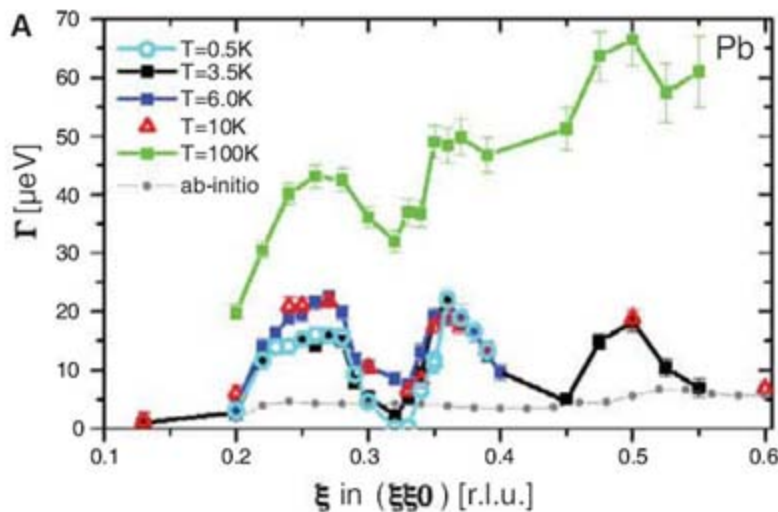
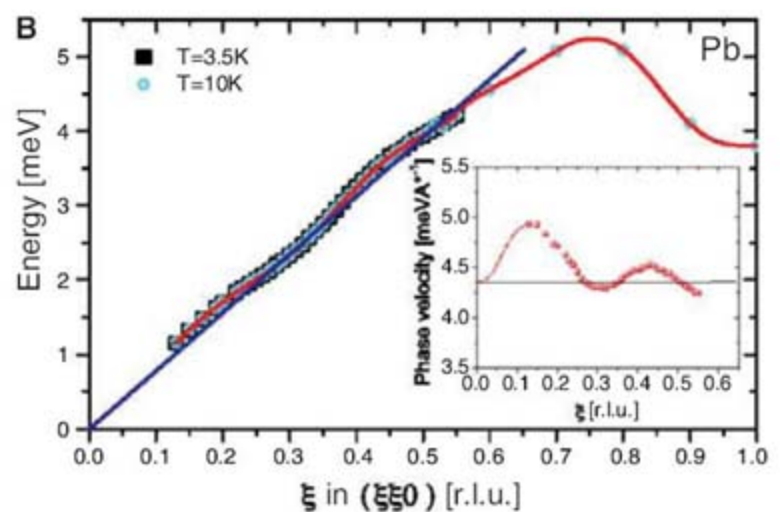


Fig. 2. (A) Linewidths of transverse acoustic phonons along $q = (\xi, \xi, 0)$ in Pb at selected temperatures. The data were obtained by correcting the measured spin-echo decay rates for instrumental effects (13). The gray symbols are the results of ab initio lattice-dynamical calculations, as described in the text. Error



bars indicate the statistical errors. (B) Dispersion relation of the same phonon extracted from triple-axis data. (Inset) The phonon phase velocity (E/q) computed from the data. The blue line in (B) and the black line in the (B) inset represent the experimentally determined sound velocity (29).

electronic density of states above $2\Delta(T)$, which moves to progressively higher energies upon cooling and closely tracks the energy gap determined in prior tunneling measurements (17) (Fig. 3 inset). Surprisingly, however, the superconductivity-induced maximum of Γ_{e-p} merges with the Kohn anomaly as $T \rightarrow 0$. At $T = 0.5$ K, both anomalies are indistinguishable within the measurement error.

In order to explore whether the coincidence of $2\Delta(T \rightarrow 0)$ and the Kohn anomaly in Pb is accidental, we performed similar experiments on phonons in Nb, an elemental superconductor with a different Fermi surface and phonon spectrum. Figure 4A shows the momentum-dependent linewidths of the transverse acoustic phonon branch along $q = (\xi, \xi, 0)$ in Nb at temperatures above and below $T_c = 9.3$ K. The data above T_c are in fair agreement with prior work (9), but they reveal

several sharp features that have not been identified before. In part on the basis of ab initio lattice dynamical calculations, they can be identified as Kohn anomalies (see below). The existence of a Kohn anomaly at $\xi \sim 0.17$ persisting up to room temperature has been suggested on the basis of prior experimental work (7, 18). As described above for Pb, the linewidths are reduced below and enhanced above the gap for quasi-particle-pair production, $2\Delta(T)$, in the superconducting state, and the low-temperature electron-phonon linewidth shows the expected dependence on wave vector (or energy). Similar to the observation in Pb, the $2\Delta(T \rightarrow 0)$ extracted from the low-temperature Γ_{e-p} of Nb again coincides with the lowest-energy Kohn anomaly within the experimental error (Fig. 4B).

To help interpret these observations, we calculated the phonon dispersions and linewidths

in the framework of ab-initio density functional perturbation theory in the local-density approximation (LDA) (13) on a very fine mesh of \mathbf{q} points in reciprocal space. The phonon frequencies were obtained by diagonalization of the dynamical matrices and the electron-phonon linewidths by Allen's formula (19). The results are in reasonable overall agreement with the experimental data (Figs. 3 and 4). In particular, both the phonon frequencies and the linewidths associated with Kohn anomalies in the high-energy transverse acoustic phonons of Nb (Fig. 4A) and in the longitudinal phonon of Pb (13) are well described, indicating that the resolution of the calculations is sufficient to reproduce subtle structures in \mathbf{q} space.

The lowest-lying Kohn anomalies in the transverse-acoustic phonon branches of both Pb and Nb are, however, not reproduced by the calculations (Figs. 3A and 4, A and B). These anomalies therefore originate in factors not included in the calculations, such as the relativistic spin-orbit coupling, phonon nonadiabaticity (20), or many-body correlations beyond Allen's formula (21, 22) or the LDA. Because the Kohn anomalies in Pb and Nb are of comparable strength, the spin-orbit coupling (which is much stronger in Pb than in Nb) cannot be responsible. Because of the large Fermi energies of both materials, nonadiabatic electron-phonon coupling effects should also be extremely weak.

This leaves electron correlation effects beyond the LDA as the most likely mechanism responsible for the low-energy Kohn anomalies. It seems reasonable to assume that the same correlations are also responsible for the observed coincidence of $2\Delta(T \rightarrow 0)$ with the same anomalies. Because the anomalies persist to temperatures above 100 K, superconducting fluctuations are unlikely to be directly responsible. We note, however, that the formation of spin or charge density waves driven by electron correlations has been predicted for Pb and other elemental metals (23). Although extensive searches for static

Fig. 3. Contribution of the electron-phonon interaction to the linewidth of the transverse acoustic phonon along $q = (\xi, \xi, 0)$ in Pb. The corresponding E is provided by the scale at the top; note the non-linear E -versus- q relationship (Fig. 2B). The lines are results of least-squares fits to the Bardeen-Cooper-Schrieffer (BCS) excitation spectrum function (30). Error bars indicate the statistical errors. (Inset) The temperature dependence of the superconducting energy gap (squares) and Kohn anomaly (triangles) extracted from the fits. The line in the inset shows the BCS expression for the superconducting gap (2), which was experimentally confirmed by tunneling spectroscopy (17). The error bars account for systematic errors associated with extracting the peak positions. At $T = 0.5$ K, for instance, fits to two peaks with energies at the end points of the error bar are in acceptable agreement with the data.

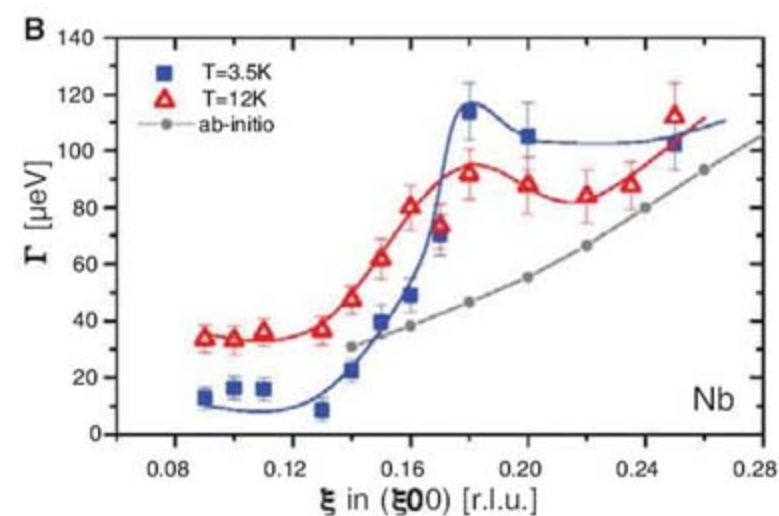
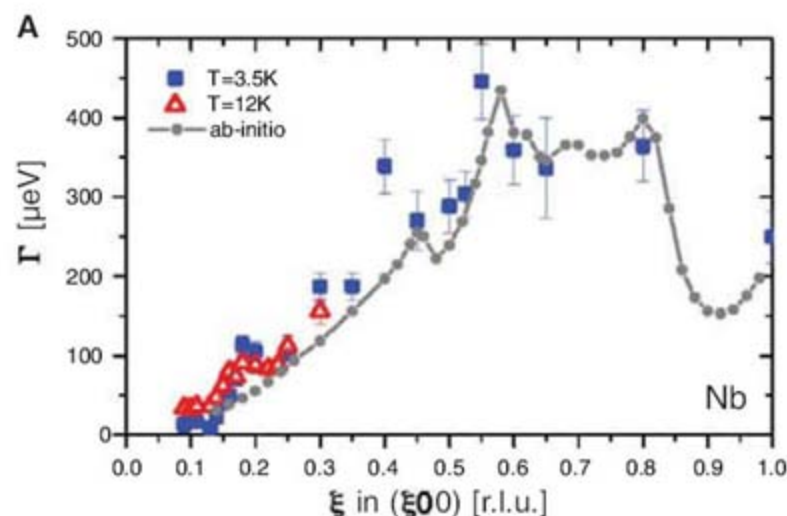
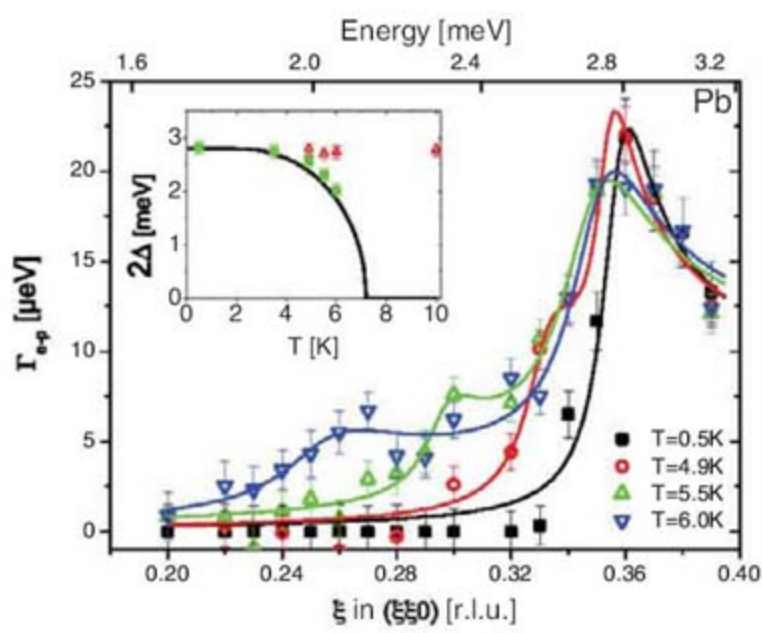


Fig. 4. (A) Linewidths of transverse acoustic phonons along $q = (\xi, 0, 0)$ in Nb at two different temperatures. The gray symbols are the results of lattice-dynamical calculations, as described in the text. (B) Blowup

of the low- q segment of (A). The corresponding E is provided by the scale at the top. The lines are guides to the eye. Error bars indicate the statistical errors.

density waves in simple metals have been unsuccessful, it is conceivable that fluctuations characteristic of such states dynamically enhance the nesting properties of the Fermi surface and hence the propensity for Kohn anomalies in the phonon spectrum. Indeed, experiments on charge-density-wave materials such as NbSe₂ have revealed Kohn anomalies (8) and Fermi-surface "pseudogaps" (24, 25) in the extended fluctuation regime at temperatures well above the onset of static density-wave order. Detailed theoretical work is required to assess whether interference between density-wave and superconducting correlations can limit the growth of the superconducting gap and lead to the observed convergence of both energy scales at low temperatures.

Our experiments on two different elemental superconductors demonstrate that the low-temperature limit of the superconducting energy gap coincides with low-lying Kohn anomalies in transverse acoustic phonons. Because both superconductors exhibit different lattice structures, phonon spectra, Fermi surfaces, and superconducting gaps, this coincidence cannot be accidental. Although its origin is presently unclear, a specific scenario to explore in future theoretical work is the interplay between density-wave and superconducting correlations. Lastly, we point out a possible analogy to research on high-temperature superconductors, where an anomalous coinci-

dence of the superconducting gap with a weakly temperature-dependent pseudogap has recently been reported in some regions of momentum space (26–28).

References and Notes

1. J. Bardeen, L. N. Cooper, J. R. Schrieffer, *Phys. Rev.* **108**, 1175 (1957).
2. For a review, see J. R. Schrieffer, *Theory of Superconductivity* (Perseus, New York, 1999).
3. For recent advances in ab initio computation of superconducting properties, see M. A. L. Marques *et al.*, *Phys. Rev. B* **72**, 024546 (2005).
4. W. Kohn, *Phys. Rev. Lett.* **2**, 393 (1959).
5. B. N. Brockhouse, T. Arase, G. Caglioti, A. D. B. Woods, *Phys. Rev.* **128**, 1099 (1962).
6. R. Stedman, L. Almquist, G. Nilsson, G. Raunio, *Phys. Rev.* **163**, 567 (1967).
7. Y. Nakagawa, A. D. B. Woods, *Phys. Rev. Lett.* **11**, 271 (1963).
8. D. E. Moncton, J. D. Axe, F. J. DiSalvo, *Phys. Rev. B* **16**, 801 (1977).
9. S. M. Shapiro, G. Shirane, J. D. Axe, *Phys. Rev. B* **12**, 4899 (1975).
10. S. P. Bayrakci, T. Keller, K. Habicht, B. Keimer, *Science* **312**, 1926 (2006).
11. T. Keller *et al.*, *Phys. Rev. Lett.* **96**, 225501 (2006).
12. T. Keller *et al.*, *Appl. Phys. A* **74**, S332 (2002).
13. Materials and methods are available on Science Online.
14. J. R. Anderson, A. V. Gould, *Phys. Rev.* **139**, A1459 (1963).
15. H. J. Maris, W. E. Massey, *Phys. Rev. Lett.* **25**, 220 (1970).
16. L. P. Pitaevski, Y. B. Levinson, *Phys. Rev. B* **14**, 263 (1976).

17. R. F. Gasparovic, B. N. Taylor, R. K. Eck, *Solid State Commun.* **4**, 59 (1966).
18. M. A. Moore, D. I. Paul, *Solid State Commun.* **9**, 1303 (1971).
19. P. B. Allen, *Phys. Rev. B* **6**, 2577 (1972).
20. S. Pisana *et al.*, *Nat. Mat.* **6**, 198 (2007).
21. E. Cappelluti, *Phys. Rev. B* **73**, 140505 (2006).
22. O. V. Dolgov, O. K. Andersen, I. I. Mazin, *Phys. Rev. B* **77**, 014517 (2008).
23. A. W. Overhauser, *Phys. Rev.* **128**, 1437 (1962).
24. S. V. Borisenko *et al.*, preprint available at www.arxiv.org/abs/0704.1544.
25. T. Kiss *et al.*, *Nat. Phys.* **3**, 721 (2007).
26. K. Tanaka *et al.*, *Science* **314**, 1910 (2006); published online 15 November 2006 (10.1126/science.1133411).
27. K. Terashima *et al.*, *Phys. Rev. Lett.* **99**, 017003 (2007).
28. W. S. Lee *et al.*, *Nature* **450**, 81 (2007).
29. D. L. Waldorf, *Bull. Am. Phys. Soc.* **5**, 170 (1960).
30. R. C. Dynes, V. Narayanamurti, J. P. Garno, *Phys. Rev. Lett.* **41**, 1509 (1978).
31. We thank O. K. Andersen, S. P. Bayrakci, S. V. Borisenko, M. Cardona, O. Dolgov, E. K. U. Gross, and G. Khalullin for fruitful discussions; C. T. Lin for the Pb crystals; J. Mayor for the Nb crystals; and K. Buchner for technical assistance.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1154115/DC1

Materials and Methods

Fig. S1

References and Notes

13 December 2007; accepted 1 February 2008

Published online 21 February 2008;

10.1126/science.1154115

Include this information when citing this paper.

Synthesis of Macrocyclic Copolymer Brushes and Their Self-Assembly into Supramolecular Tubes

Michel Schappacher and Alain Deffieux*

We report on an efficient route to design large macrocyclic polymers of controlled molar mass and narrow dispersity. The strategy is based on the synthesis of a triblock copolymer ABC, in which the long central block B is extended by two short A and C sequences bearing reactive antagonist functions. When reacted under highly dilute conditions, this precursor produces the corresponding macrocycle by intramolecular coupling of the A and C blocks. Chloroethyl vinyl ether was selected as the monomer for the central block B, because it can be readily derivatized into brushlike polymers by a grafting process. The corresponding macrocyclic brushes were decorated with polystyrene or randomly distributed polystyrene and polyisoprene branches. In a selective solvent for the polyisoprene branches, the macrocyclic brushes self-assemble into cylindrical tubes of up to 700 nanometers.

The interest in cyclic macromolecules began more than 50 years ago with the theoretical prediction of the effect of chain cyclization on polymer properties (1, 2) and the discovery of naturally occurring macrocycles such

as DNA (3). Recently, the self-assembly of proteins into nanorings for the fabrication of advanced materials has been reported (4). Because of their very limited availability, large polymer macrocycles remain a fascinating curiosity for theoreticians and physicists and a challenging area for synthetic chemists (5).

Macrocyclic polymers were first obtained from macromolecules systems exhibiting ring-linear chain equilibria (6). In such systems, low-to-medium molar mass macrocycles are obtained

generally in admixture with linear chains, although a more selective approach has been recently reported by Grubbs (7).

The most appropriate method for the synthesis of cyclic polymers with controlled size and narrow dispersity was first proposed by Casassa (8) more than 40 years ago. It is based on the end-to-end chain coupling of linear α,ω -difunctional chains in highly dilute conditions. Coupling of α,ω -dianionic polymers with difunctional agents has been the most extensively used ring closure approach (9–11); however, cyclization yields for large macrocycles are generally low, and fractionation procedures are required to remove residual linear chains and polycondensates (12). An alternative route is based on the end-to-end coupling of α,ω -heterodifunctional linear chain (13). Cyclization is performed under high dilution by selective activation of one polymer end, which reacts intramolecularly with the second chain end. Higher cyclization yields have been reported but molar masses of the macrocycles remain limited. A more selective approach was proposed by Tezuka (14) that involves the pre-cyclization of linear chains bearing ionic end groups. However, only very low molar mass macrocycles have been synthesized.

The preparation of large macrocyclic polymers and copolymers is thus limited (12) by (i) the difficulty to get pure α,ω -difunctional high molar mass precursors, (ii) the drastic decrease of the end-to-end ring closing efficiency when in-

CNRS, Université Bordeaux 1, Ecole Nationale Supérieure de Chimie et Physique de Bordeaux (ENSCP), Laboratoire de Chimie des Polymères Organiques, Unité Mixte de Recherche 5629, 16 Avenue Pey Berland, 33607 Pessac cedex, France.

*To whom correspondence should be addressed. E-mail: deffieux@enscpb.fr

creasing the distance between chain ends, and (iii) the problem of elimination of linear contaminants of the same high molar mass.

We have developed a method to synthesize large polymer macrocycles based on the synthesis of an ABC triblock copolymer, in which the long central B block is extended by two short A and C sequences bearing monomer units with reactive antagonist functions (Fig. 1). The external blocks are then selectively activated under dilute conditions to allow intramolecular coupling between the A and C blocks and form the macrocyclic polymers, with further functionalization to form the corresponding brush polymers with polystyrene (PS) or randomly distributed polystyrene and polyisoprene (PS/PI) branches.

The ABC copolymers were obtained by sequential living cationic polymerization of three different vinyl ethers (fig. S1). The short A block consists of units of *tert*-butyl dimethyl silyl ethoxy vinyl ether (a hydroxy-protected vinyl ether). Chloroethyl vinyl ether (CEVE) was selected as the second monomer for the B block, because PCEVE [poly(CEVE)] can be readily converted into the corresponding PS brush structure (15). Finally, block C was grown by addition of cyclohexane dimethanol divinyl ether (a divinyl ether), which allows introducing monomer units bearing vinyl ether side groups in appropriate reaction conditions. Detailed information on the synthesis methods can be found in the supporting online material.

Typically, the triblock copolymers were constituted to A, B, and C blocks with targeted average degree of polymerization (\overline{DP}_n) of roughly 5, 1000, and 5, respectively. The properties of the ABC copolymer "(L)" are given in Table 1.

After deprotection of the hydroxyls on the block A's, cyclization was achieved under high dilution via their intramolecular reaction with the vinyl ether functions on block C's in the presence of an acid salt as a catalyst. The size exclusion chromatograms (SECs) of the ABC copolymer before and after cyclization are shown in Fig. 2.

The shift of the main SEC peak to higher elution volume is consistent with the reduction of the hydrodynamic volume of the macromolecule after chain cyclization (16). The slight peak broadening and the presence of two small peak shoulders indicate the presence of a limited amount of residual linear precursor and of linear or cyclic dimers. Some characteristics of cyclic copolymer "(C)" that formed are indicated in Table 1. It is difficult to get more precise and direct information on the architecture of the polymers formed and on the cyclization yield, with conventional analytical techniques.

The synthesis of PS brushes by grafting living polystyryl lithium (PSLi) and/or polyisoprenyl lithium (PILi) onto PCEVE chains has been studied in our group (17, 18). One advantage of this approach relies on the capacity to precisely control the architecture and dimensions of the brush, given that both the PCEVE backbone and the PS grafts are obtained by living polymerizations. In addition,

the good efficiency and selectivity of the coupling reaction allow the preparation of brush polymers exhibiting the same architecture as the initial PCEVE backbone as shown by atomic force microscopy (AFM) (19). This strategy was applied here to convert large macrocyclic polymers into the corresponding cyclic PS brushes. This approach had been tried with cyclic PCEVE of low molar mass prepared from an α,ω -heterodifunctional precursor (20), but the formation of cyclic brush was not demonstrated.

The SEC chromatogram of a cyclic brush-Ca is shown in Fig. 2. The main elution peak can be attributed to macrocyclic brushes, whereas the two small shoulders observed at lower elution volume have been assigned on the basis of static light scattering (SLS) measurements (21) to linear PS brushes and to brush dimers of various architec-

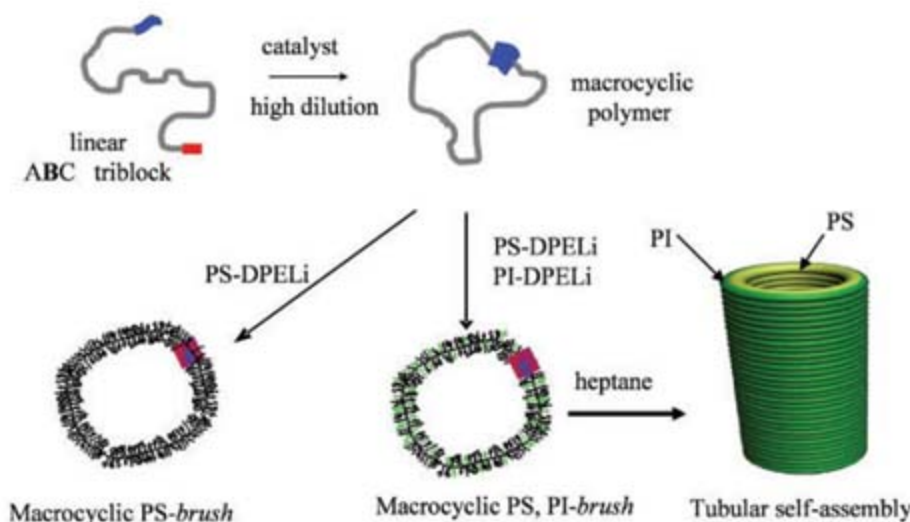


Fig. 1. Strategy for the synthesis of cyclic comblike copolymers and their self-assembly. PS-DPELi, (1,1-diphenylethylene) end-capped polystyryllithium; PI-DPELi, (1,1-diphenylethylene) end-capped polyisoprenyllithium.

Table 1. Dimensional characteristics of linear (L) and cyclized (C) ABC copolymer and of the corresponding macrocyclic brushes with different PS or PS/PI branch lengths. M_n , number-average molecular weight; M_w , weight-average molecular weight; M_n^{app} , apparent M_n ; M_p^{app} , apparent SEC peak molecular weight.

Chain architecture*	\overline{DP}_n^\dagger	$M_n^{app}\ddagger$ ($\times 10^{-3}$)	M_w/M_n	$M_p^{app}\ddagger$ ($\times 10^{-3}$)	M_w^\S ($\times 10^{-3}$)
L	0	92	1.05	99	95
C	0	78	1.15	76	96
PS brush-Ca	90/0	740	1.14	800	5900
PS brush-Cb	260/0	1650	1.09	2160	10,200
PS/PI brush-Cc	170/50	872	1.10	967	6800

* \overline{DP}_n of PCEVE is 870. $\dagger\overline{DP}_n$ of PS or PI branches. \ddagger Determined by SEC in THF by means of PS standards. \S Measured by SLS with refractive index increments (dn/dc) of 0.09 (PCEVE L and C), 0.18 (PS brushes), and 0.15 (PS/PI brush). $||$ Number of PS and PI branches per macromolecule are 400 and 350, respectively, which corresponds to 80% in weight of PS.

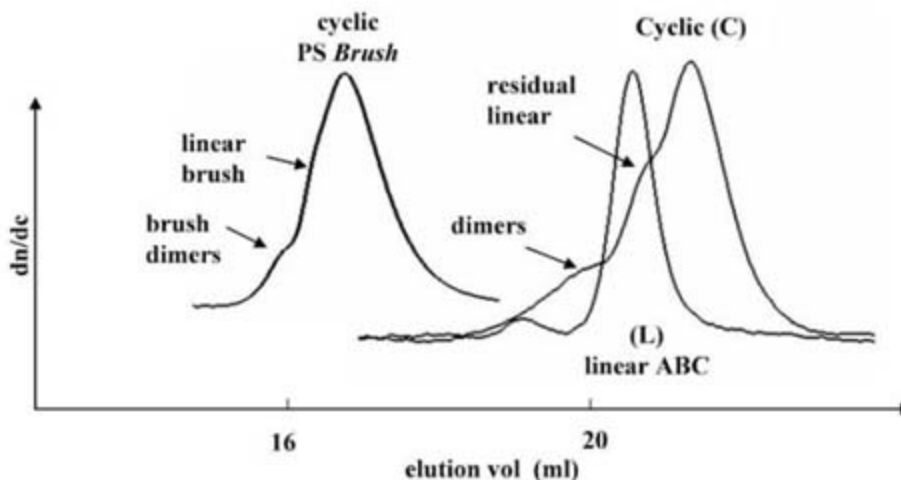


Fig. 2. SECs in tetrahydrofuran (THF) of ABC copolymer (L), cyclized polymer (C), and corresponding cyclic PS brush-Ca (see Table 1).

tures. Direct imaging (19, 22–24) allowed us to confirm these conclusions and to get direct and quantitative information on the cyclic polymer structures obtained. The AFM topographic image of *brush-Ca* (Fig. 3A) shows single macromolecules isolated on the highly oriented pyrolytic graphite surface. Most of the structures observed are monodispersed macrocyclic brushes, which agrees with a cyclization efficiency of 50 to 60%. Among the other polymer architectures observed are linear brushes corresponding to uncyclized ABC copolymer, linear and macrocyclic dimers exhibiting about twice the size of the linear precursor or of the corresponding cyclics, and tadpole structures.

Samples containing larger fractions of macrocyclic brushes were readily obtained by selective precipitation from cyclohexane solution upon gradually decreasing the temperature from 19° to 18°C, corresponding to the theta conditions of comb PS. The AFM image of the fractionated *brush-Ca* is shown in Fig. 3B. Typically, the fractionated samples contain about 80 to 85% of cyclic brushes. The main characteristics of the cyclic brushes are provided in Table 1. The influence of the PS graft dimensions on the characteristics of the brush macrocycles is shown in Fig. 3, C and D, for samples with (\overline{DP}_n) of 90 and 260, respectively. The two cyclic brushes obtained from a PCEVE with a \overline{DP}_n of 870 are characterized by a mean average diameter of 65 to 70 nm for the central backbone (fig. S2). This yields a circumference of 205 to 220 nm for the cyclized PCEVE. Assuming a \overline{DP}_n of 870 and a value of 2.5 Å per monomer unit, this corresponds to a fully extended backbone conformation (870 nm \times 0.25 nm = 217.5 nm). This

can be explained by the steric hindrance and compactness near the PCEVE backbone resulting from the high PS grafting density, which according to the brush molar mass (Table 1) corresponds to about three PS grafts every four CEVE units. However, no chain breaking was observed despite the strong mechanical forces applied to the PCEVE backbone. The crown section and the size of the central cavity are determined by the branch length and can be tuned by controlling the PS \overline{DP}_n . As it may be seen (Fig. 3, A and D), the macrocycles as well as other brush architectures are separated by an almost constant distance, which corresponds to the length of two PS branches in an almost fully extended conformation (fig. S2).

To obtain the corresponding cyclic copolymer brushes with PS/PI branches, we successively reacted polystyryllithium and polyisoprenyllithium both end-capped with 1,1-diphenylethylene in determined amounts with the macrocyclic PCEVE to yield cyclic brushes with PS/PI branches that were randomly distributed (Fig. 1). The characteristics of one cyclic PS/PI brush used in this study are indicated in Table 1. Its capacity to self-organize was investigated in heptane—a selective solvent of the PI blocks—by dynamic light scattering (DLS) and by AFM as solid deposits from their solutions. In heptane, the cyclic PS/PI brush copolymer self-assembles, forming a single and monodisperse population of large cylindrical objects with an average hydrodynamic radius of 200 nm. The corresponding DLS data are available as supporting online material (fig. S3). Further insight in these supramolecular structures was gained by AFM analysis (Fig. 3, E and F). The selective formation of rigid tubular objects (with section

diameters of about 100 nm, corresponding to elementary macrocyclic PS/PI brushes, and lengths of up to 700 nm) that were either isolated or aggregated is observed. This agrees with the self-assembly of elementary cyclic brushes directed by intermolecular association between PS branches, whereas the PI grafts ensure their dispersion/solubilization into the selective PI solvent. Although they are not observed in heptane solution, the formation of clusters of tubular objects observed on AFM images (Fig. 3E) likely takes place during solvent evaporation on graphite. The length of the supramolecular tubes, as well as the observed structural defects, results from the contribution of remaining linear and other brush structures that perturb or stop the self-assembly of brush macrocycles.

We have developed a selective route to large macrocyclic polymers based on the coupling of external short blocks of ABC-type copolymers. The substantial increase in cyclization efficiency can be attributed both to the high functionality of ABC precursor and to the presence of several reactive functions at each chain end that increase the probability of intramolecular end-to-end cyclization. As indicated by the absence of cross-linked chains or high polycondensates, the presence of several reactive groups at each chain end does not yield significant multichain coupling under high dilution.

The obtained macrocyclic brushes constitute a class of macromolecules with a precise and tunable chain architecture that can be used as such or as elementary building blocks for the preparation of supramolecular tubes, the structure of which can be eventually frozen by cross-linking. By changing the character and proportion of the constitutive blocks, tubular assemblies are expected

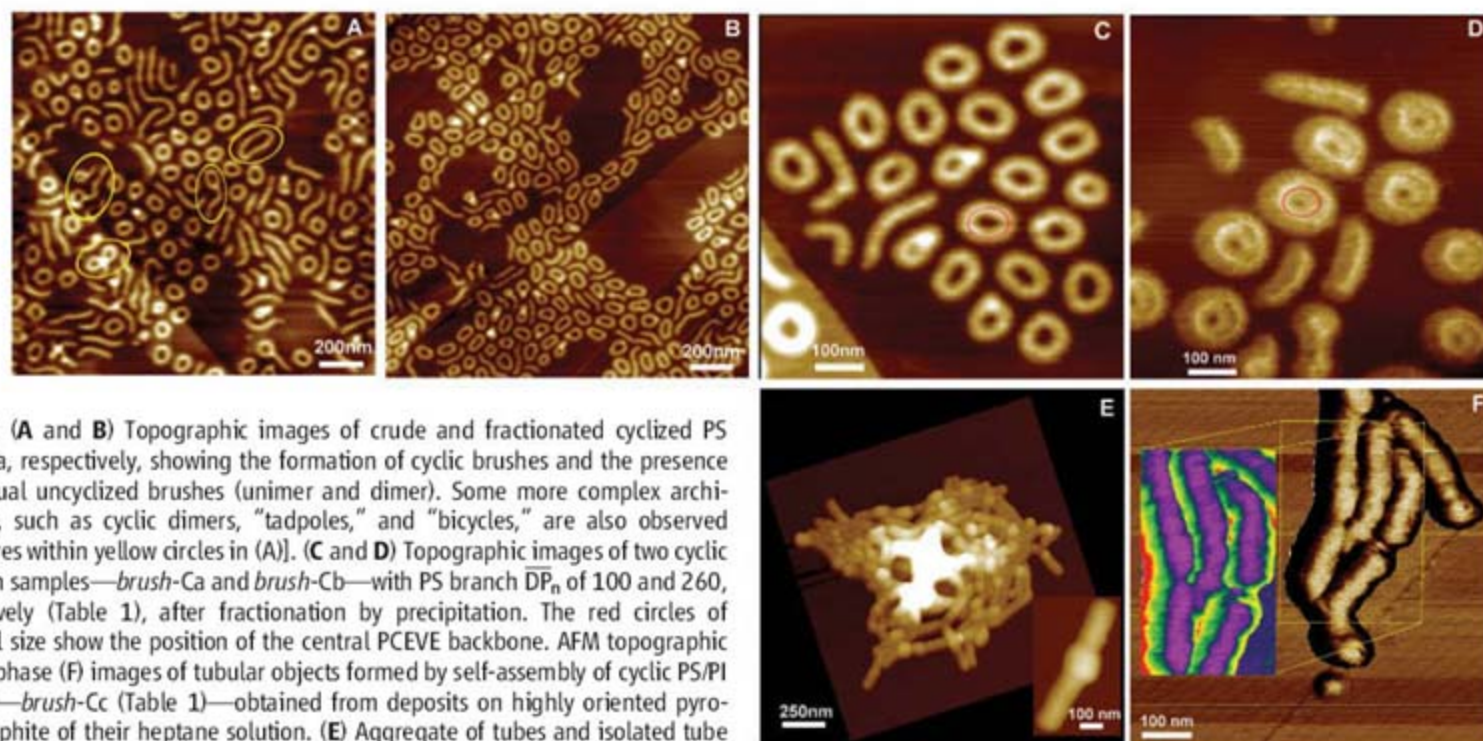


Fig. 3. (A and B) Topographic images of crude and fractionated cyclized PS *brush-Ca*, respectively, showing the formation of cyclic brushes and the presence of residual uncyclized brushes (unimer and dimer). Some more complex architectures, such as cyclic dimers, “tadpoles,” and “bicycles,” are also observed [structures within yellow circles in (A)]. (C and D) Topographic images of two cyclic PS brush samples—*brush-Ca* and *brush-Cb*—with PS branch \overline{DP}_n of 100 and 260, respectively (Table 1), after fractionation by precipitation. The red circles of identical size show the position of the central PCEVE backbone. AFM topographic (E) and phase (F) images of tubular objects formed by self-assembly of cyclic PS/PI brushes—*brush-Cc* (Table 1)—obtained from deposits on highly oriented pyrolytic graphite of their heptane solution. (E) Aggregate of tubes and isolated tube (diameter, 100 nm; length, 600 nm). (F) Series of tubes interconnected by their polyisoprene shell (black) and image in reverse mode showing the internal PS (purple) and external PI (green) parts. The stripes corresponding to the elementary macrocyclic copolymer brushes are also visible.

to be formed in different media, whereas the inner surface of the tube can be tuned for a targeted application by selecting a second type of constitutive block of appropriate characteristics and properties.

References and Notes

- B. H. Zimm, W. H. Stockmayer, *J. Chem. Phys.* **17**, 1301 (1949).
- V. Bloomfield, B. H. Zimm, *J. Chem. Phys.* **44**, 315 (1966).
- R. Dulbecco, M. Vogt, *Proc. Natl. Acad. Sci. U.S.A.* **50**, 236 (1963).
- J. C. T. Carlson *et al.*, *J. Am. Chem. Soc.* **128**, 7630 (2006).
- A. Deffieux, R. Borsali, in *Macromolecular Engineering: Precise Synthesis, Materials Properties, Applications*, K. Matyjaszewski, Y. Gnanou, L. Leibler, Eds. (Wiley-VCH, Weinheim, Germany, vol. 2, 2007), pp. 875–908.
- J. A. Semlyen, G. R. Walker, *Polymer* **10**, 597 (1969).
- C. W. Bielawski, D. Benitz, R. H. Grubbs, *Science* **297**, 2041 (2002).
- E. F. Cassassa, *J. Polym. Sci. A* **3**, 605 (1965).
- B. Vollmert, J. X. Huang, *Makromol. Chem. Rapid Commun.* **2**, 467 (1981).
- G. Hild, C. Strazielle, P. Rempp, *Eur. Polym. J.* **19**, 721 (1983).
- J. Roovers, P. Toporowski, *Macromolecules* **16**, 843 (1983).
- T. E. Hogen-Esch, *J. Polym. Sci. Polym. Chem.* **44**, 2139 (2006).
- M. Schappacher, A. Deffieux, *Makromol. Chem. Rapid Commun.* **12**, 447 (1991).
- Y. Tezuka, H. Oike, *J. Am. Chem. Soc.* **123**, 11570 (2001).
- A. Deffieux, M. Schappacher, *Macromolecules* **32**, 1797 (1999).
- G. Hadziioannou *et al.*, *Macromolecules* **20**, 493 (1987).
- M. Schappacher, A. Deffieux, *Macromolecules* **33**, 7371 (2000).
- M. Schappacher, Z. Mughtar, A. Deffieux, *Macromolecules* **34**, 7595 (2001).
- P. Viville *et al.*, *Polymer* **45**, 1833 (2004).
- M. Schappacher, C. Billaud, C. Paulo, A. Deffieux, *Macromol. Chem. Phys.* **200**, 2377 (1999).
- Assignment of peaks and peak shoulders on the SEC of polymer brushes to dimer, linear, and cyclic structures (Fig. 2) was made on the basis of their radius of gyration measured by SLS with an online laser light scattering detector.
- M. Schappacher, A. Deffieux, *Macromolecules* **38**, 4942 (2005).
- S. S. Sheiko, M. Möller, *Chem. Rev.* **101**, 4099 (2001).
- S. S. Sheiko, in *Advances in Polymer Science: New Developments in Polymer Analytics II* (Springer Berlin, Heidelberg, vol. 151, 2000), pp. 61–174.

Supporting Online Material

www.sciencemag.org/cgi/content/full/319/5869/1512/DC1

Materials and Methods

Figs. S1 to S3

References

6 December 2007; accepted 1 February 2008

10.1126/science.1153848

Splitting of the 520-Kilometer Seismic Discontinuity and Chemical Heterogeneity in the Mantle

Ashima Saikia,* Daniel J. Frost, David. C. Rubie

Seismic studies indicate that beneath some regions the 520-kilometer seismic discontinuity in Earth's mantle splits into two separate discontinuities (at ~500 kilometers and ~560 kilometers). The discontinuity near 500 kilometers is most likely caused by the (Mg,Fe)₂SiO₄ β-to-γ phase transformation. We show that the formation of CaSiO₃ perovskite from garnet can cause the deeper discontinuity, and by determining the temperature dependence for this reaction we demonstrate that regional variations in splitting of the discontinuity arise from variability in the calcium concentration of the mantle rather than from temperature changes. This discontinuity therefore is sensitive to large-scale chemical heterogeneity. Its occurrence and variability yield regional information on the fertility of the mantle or the proportion of recycled oceanic crust.

Recent regional seismic observations have identified multiple discontinuities near the depth of the 520-km discontinuity in the transition zone (1–4). Beneath some regions, this discontinuity splits into two discontinuities, one at a depth of about 500 km and the other near 560 km (1). The depth of these two discontinuities varies, implying that they are marking phase transitions that are responding to variations in composition between regions of the mantle or to varying temperature. The wadsleyite-to-ringwoodite (β-γ) transition is often implicated as the cause of the 520-km discontinuity (5). However, the exsolution of CaSiO₃ perovskite from majoritic garnet that also occurs at a similar depth could cause the observed higher pressure split in the 520-km discontinuity (6). This reaction is complex because majoritic garnet is a multicomponent solid solu-

tion. Existing data (7–10) are not sufficiently consistent for modeling how the reaction varies with respect to pressure, temperature, and mantle composition. We studied the exsolution of Ca perovskite in high-pressure and -temperature experiments in order to ascertain the depth interval over which the reaction occurs. We also estimated the jump in density and sound velocity expected for this reaction and used these data to explain regional occurrences of single and split 520-km discontinuities.

At mid-transition-zone pressures (17 GPa), garnet will contain all of the Ca in most plausible mantle chemical compositions and will have a substantial majorite component (i.e., Mg₄Si₄O₁₂), which results from the substitution of Si and Mg into the Al site at high pressure (11, 12). The proportions of Ca and Al in transition-zone garnets will vary strongly between different rock types, being low in rocks of peridotite composition but much higher in rocks that are remnants of subducted oceanic crust, that is, with the composition of midocean ridge basalt (MORB). With increasing pressure, Ca perovskite exsolves from

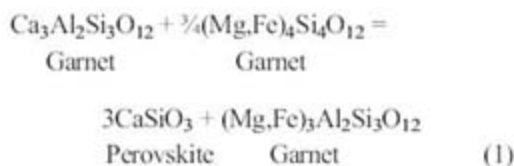
garnet as the solubility of CaSiO₃ in garnet decreases. We measured this solubility in multianvil experiments between 15 and 24 GPa at 1400° and 1600°C and as a function of the garnet majorite content (i.e., Al/Si ratio). Ca-free glass compositions on the join (Mg,Fe)₄Si₄O₁₂–(Mg,Fe)₃Al₂Si₃O₁₂ were mixed with presynthesized CaSiO₃ wollastonite in a 1:2 weight ratio to produce starting materials for multianvil experiments. These samples were loaded into rhenium capsules that contained four sample chambers (Fig. 1). Three samples with varying Al/Si ratios were loaded into each experiment, plus a (Mg,Fe)₂SiO₄ sample to calibrate the pressure. During the experiment, the glass and CaSiO₃ compositions crystallize rapidly to form garnet and perovskite, respectively, and CaSiO₃ dissolves into the garnet (13). The diffusion into the garnet is slow, and experiments must be equilibrated for at least 24 hours in order to reach equilibrium. A B₂O₃ flux was also added to speed up the reaction. Some experiments were reversed by first synthesizing Ca-bearing garnets and measuring their compositions after CaSiO₃ had subsequently exsolved at a particular pressure and temperature. After the experiments had been quenched, recovered samples were analyzed with an electron microprobe. During the experiments, the (Mg,Fe)₂SiO₄ olivine pressure calibrant crystallized to form coexisting high-pressure phases. The pressure could be determined because the high-pressure Mg₂SiO₄-Fe₂SiO₄ phase diagram consists of a series of two-phase divariant loops where the Fe/(Fe+Mg) ratios of coexisting phases are very sensitive to pressure. In this way, we calibrated accurately the pressure of Ca-perovskite formation relative to the pressure of phase transformations in the Mg₂SiO₄-Fe₂SiO₄ system (13–15).

Our data (Fig. 2) show that the solubility of CaSiO₃ in garnet in equilibrium with Ca perovskite decreases strongly with pressure, but it increases with the garnet Al/Si ratio at a given pressure, which can be seen by comparing results from the different starting compositions. At lower temperatures (1400°C), CaSiO₃ solubility in gar-

Bayerisches Geoinstitut, University of Bayreuth, D-95444 Bayreuth, Germany.

*To whom correspondence should be addressed. E-mail: ashima.saikia@uni-bayreuth.de

net also decreases. The exsolution can be described by the equation



We fit the experimental results to a thermodynamic model (13, 16) for Eq. 1 that allows us to calculate the extent of exsolution for any bulk

composition and reliably extrapolate the results to lower temperatures, where slow kinetics inhibit the achievement of equilibrium in experiments.

In the mid-transition zone, the Ca content of garnet in a fertile mantle peridotite composition (18) is constant with pressure up to about 19.7 GPa at 1600°C because garnet contains all the Ca in the bulk composition (Fig. 2A). Above this pressure, the garnet CaSiO₃ solubility limit is reached, and Ca perovskite exsolves. Because the Ca content of subducted oceanic crust (i.e., MORB) is higher, the garnet reaches saturation at lower pres-

ures than in the peridotite composition (Fig. 2B), but the exsolution of Ca perovskite is offset somewhat because the MORB garnet has a higher Al content, which shifts CaSiO₃ saturation to higher pressures. The garnet becomes more Al-rich as CaSiO₃ exsolves and, for a MORB composition, the Al/Si ratio reaches its maximum, that is, that of pyrope (i.e., Al/Si = 2/3), above 23 GPa. At this point, if Ca perovskite is to continue to exsolve, an additional phase needs to form to accommodate the excess Al from the breakdown of garnet. In some previous experiments, a

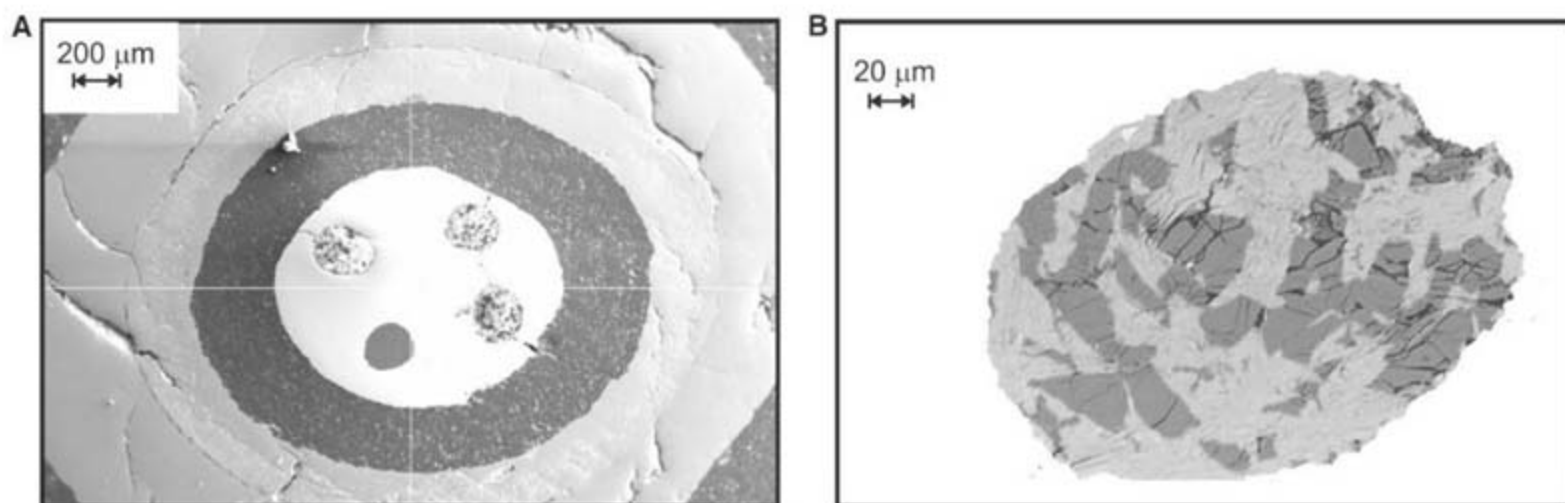


Fig. 1. Back-scattered electron images of sample capsules. (A) A radial section through the high-pressure assembly shows the outer pressure medium and furnace with an inner four-chamber Re capsule (white). Three sample chambers contain garnet and Ca-perovskite assemblages (which appear lighter), whereas

the bottom left sample chamber contains a (Mg,Fe)₂SiO₄ pressure-calibrant sample (darker). (B) An enlargement of one of the sample chambers showing darker, angular garnet grains coexisting with lighter Ca perovskite. During the experiment, CaSiO₃ diffuses into the garnet until it becomes saturated in this component.

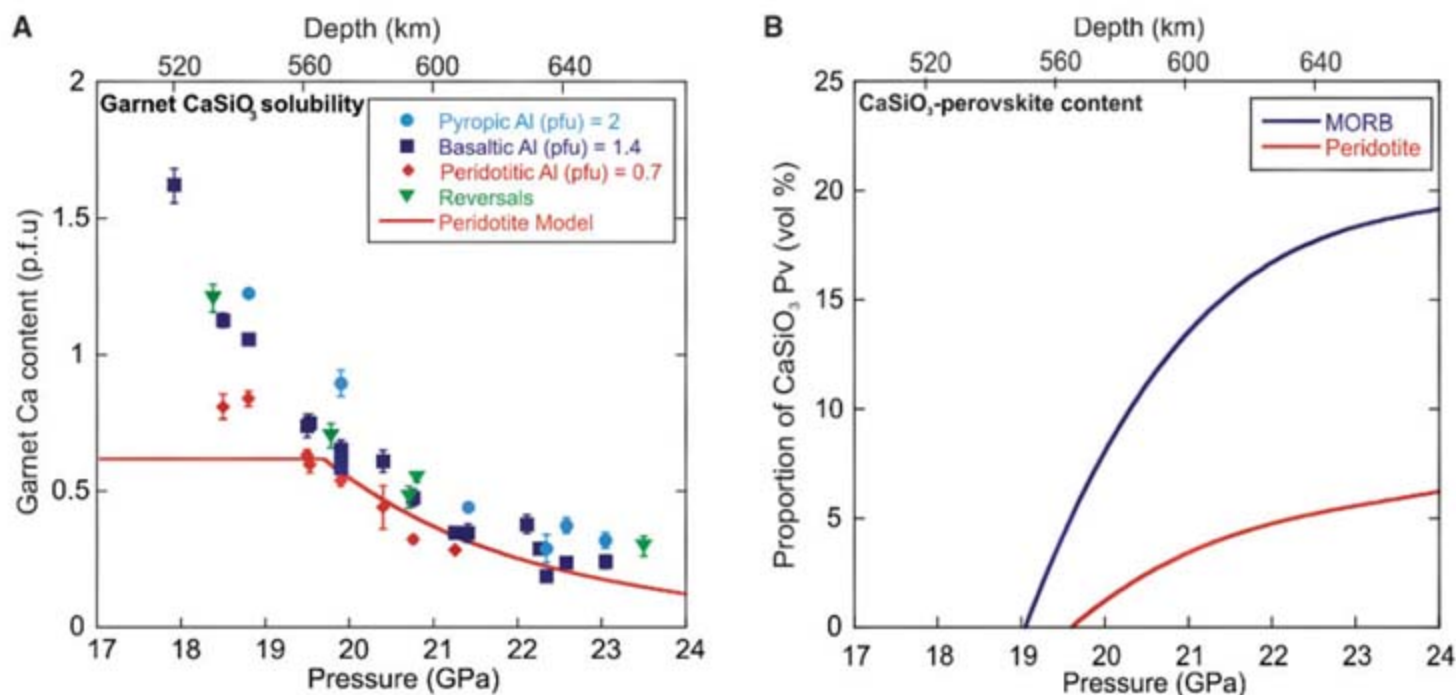


Fig. 2. (A) Experimental results showing the Ca content of garnet saturated with Ca perovskite as a function of pressure at 1600°C. Symbols refer to starting compositions, with different garnet majorite proportions labeled according to the rock types in which the garnets would be present. Al (pfu) refers to the Al content per garnet formula unit, that is, Al (pfu) = 2 refers to the formula (Mg,Fe)₃Al₂Si₃O₁₂ (i.e., pyropic). These Al contents only refer to those in the starting material because Al contents decrease as the garnets become

CaSiO₃ saturated. The reversals were performed by using presynthesized Ca-bearing garnets. The curve shows the calculated Ca content of garnet for a peridotite composition (18). Below 19.7 GPa, the Ca content is constant with pressure because the garnet contains the entire Ca content of the rock, whereas in the experiments excess CaSiO₃ was always present. (B) The volumetric proportion of CaSiO₃ perovskite that forms as a function of pressure at 1600°C in rocks of peridotite and MORB compositions, respectively.

Na- and Al-rich phase [with the general formula $\text{Na}(\text{Mg,Fe})_2\text{Al}_5\text{SiO}_{12}$] referred to as NAL has been observed that apparently serves this purpose [e.g., (8)]. The appearance of this Al-rich phase in a MORB composition is, therefore, directly coupled to the Ca perovskite-forming reaction and its influence on the Al content of the garnet.

In Earth's mantle, the exsolution of Ca perovskite is inherently nonlinear because the activity of the garnet Ca component has an exponential relationship to pressure (13, 16). Although Ca perovskite forms over an extensive depth range [about 60 km (10)], the transformation will appear sharper to seismic waves (19, 20) because a substantial portion of the exsolution occurs initially over a relatively narrow depth interval, causing a moderately sharp change in density. Combining the experimental results with mineral physics data (13, 21–23), we estimated the density and sound velocity changes that this exsolution reaction would produce (Fig. 3). For fertile peridotite at 1400°C, the β - γ transition produces a relatively strong discontinuity at 500 to 520 km, whereas the formation of Ca perovskite creates a break in slope at about 540 km, followed by a region with a higher velocity gradient, which curves back toward the ambient transition zone gradient over a depth interval of about 60 km. A similar situation occurs at 1600°C, but because the Clapeyron slope of the β - γ transition is steeper [$-0.007 \text{ GPa K}^{-1}$ (24)] than that of the Ca-perovskite reaction ($-0.004 \text{ GPa K}^{-1}$), the two discontinuities merge at 540 to 560 km. Both discontinuities contribute to a generally steep gradient in seismic velocity in the lower part of the transition zone, which is a feature observed in most seismic reference models (25). For a

MORB composition, no β - γ transition occurs, but because of the higher Ca content a strong discontinuity is caused as Ca perovskite forms.

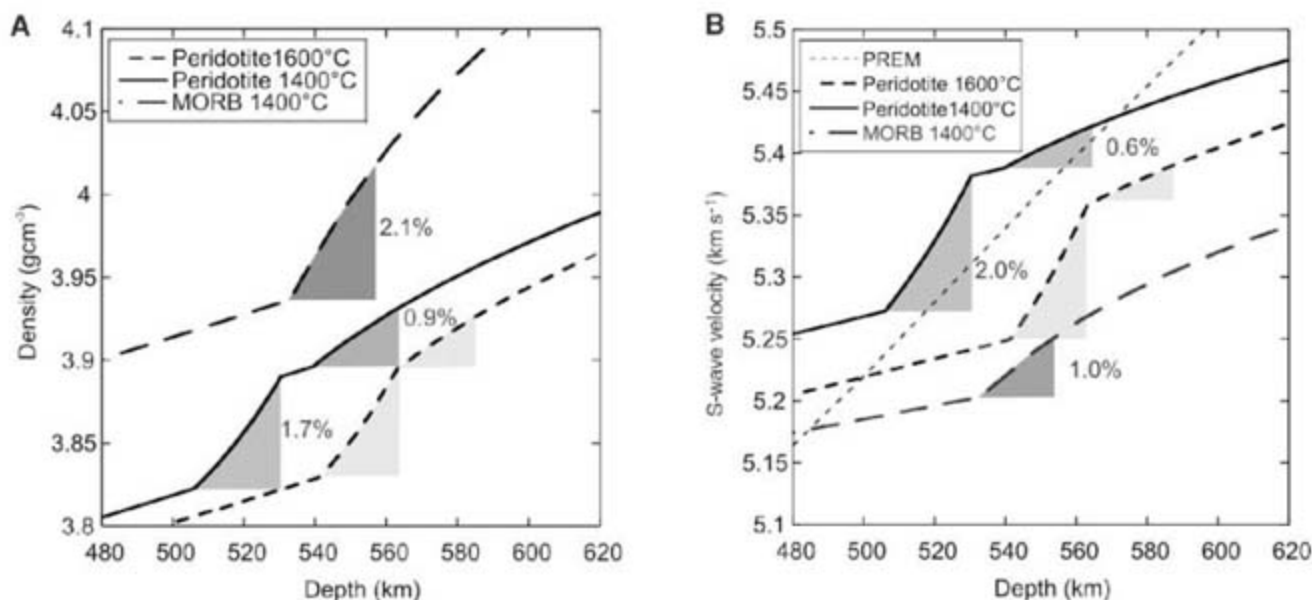
Most observations of the 520-km discontinuity have been made with studies of long-period underside SS reflections (1, 5, 26), where the amplitude (i.e., the visibility of the reflected waves) depends on the S impedance contrast [$\Delta(\rho V_s)$]. The β - γ transition occurs over a depth interval of about 25 km with an estimated $\Delta(\rho V_s)$ of 3.6%. Although the seismically observable effective depth interval of the Ca-perovskite reaction is harder to determine and the reflectivity will vary with the frequency of the seismic waves (19, 20), we can make a simple comparison by calculating $\Delta(\rho V_s)$ over the same depth interval as the β - γ transition, which gives an impedance contrast of about 1.5%. These impedance contrast estimates are in good agreement with those that have been proposed for the two discontinuities in regions where the 520 is split (1).

Many of the characteristics of splitting of the 520 could result from variations in mantle temperature. At $\sim 1600^\circ\text{C}$, both transformations would merge into a single discontinuity, as observed in certain regions of Earth, whereas at lower temperatures (1400°C) two discontinuities would be observed. Although this may explain some of the observations, it is not completely consistent with the general behavior, which is that a discontinuity at about 520 to 530 km splits in some regions into a shallower (500 km) and a deeper (560 km) discontinuity (1). If this is a result of temperature variation alone, it would imply that the Clapeyron slope of the Ca-perovskite reaction is negative, whereas both the β - γ and Ca-perovskite reactions have positive slopes. In addition, at conditions where the two discontinuities merge,

the single discontinuity should be substantially deeper than its typical depth of about 520 km (27). A more likely explanation is that, in many regions where only a single 520-km discontinuity is observed, the Ca-perovskite reaction is invisible because the Ca content of the mantle is too low. A low Ca content would result when a region has undergone partial melting near Earth's surface and thus became depleted in its fertile component. A strong 560 reflection, on the other hand, indicates the presence of fertile mantle or mantle that contains a substantial component of recycled MORB crust. The impedance contrast for the Ca-perovskite reaction in MORB is about 3%. Because the β - γ transition does not occur in a MORB composition, mantle enriched in subducted MORB crust would have a weak 500-km discontinuity but a much stronger reflection at about 560 km, which is consistent with observations (1).

One explanation for splitting of the 520 would be that it occurs in regions containing substantial amounts of subducted oceanic crust, which possesses higher Ca concentrations than normal mantle. This would be consistent with evidence for cooler-than-average mantle in these regions because the β - γ transition appears to be also slightly uplifted (1, 2) as the 520 splits, although there are substantial uncertainties in absolute discontinuity depth estimates. If subducting slabs descend at a high angle directly into the lower mantle, such regions would not be expected to be globally extensive. However, the widespread splitting of 520 would be easily explained if subducted material is accumulating at the base of the transition zone (28, 29), creating relatively low-temperature flat-lying regions rich in subducted crust. This hypothesis is supported by the geographical dis-

Fig. 3. (A) Density and (B) shear wave velocity estimated along isotherms for fertile peridotite (18) and MORB compositions as a function of depth, illustrating discontinuity structure in the region of 520 km. The phase proportions are calculated by using a thermodynamic model fitted to the experimental data in addition to auxiliary thermodynamic data for the $(\text{Mg,Fe})_2\text{SiO}_4$ β - γ phase transformation (13, 14). For the peridotite composition, two discontinuities (shaded) can be observed that result from the shallower β -to- γ and the deeper garnet exsolution of Ca-perovskite reactions. For the MORB composition, only the Ca-perovskite reaction occurs. Because the Ca-perovskite reaction is nonlinear and its seismically observable depth interval is difficult to assess, the shaded region is shown with the same depth interval as the β - γ transition for comparison. The percent density



and S-wave velocity jumps resulting from both the β - γ and Ca-perovskite reactions are shown calculated over the shaded regions. Experimental uncertainties in discontinuity depths are about 10 km; density and velocity uncertainties are discussed in (13). Preliminary Reference Earth Model (PREM) S-wave velocity is shown for comparison.

tribution of split 520 observations, which in many instances occur beneath continental regions, such as North America and Eastern Asia (1–4), under which extensive subduction has taken place. Splitting beneath the Indian Ocean and Africa, on the other hand, may be caused by the presence of particularly fertile mantle.

References and Notes

1. A. Deuss, J. Woodhouse, *Science* **294**, 354 (2001).
2. H. J. Gilbert, A. F. Sheehan, K. G. Dueker, P. Molnar, *J. Geophys. Res.* **108**, 10.1029/2001JB001194 (2003).
3. A. Deuss, S. A. T. Redfern, K. Chambers, J. H. Woodhouse, *Science* **311**, 198 (2006).
4. M. van der Meijde, S. Van der Lee, D. Giardini, *Phys. Earth Planet. Inter.* **148**, 233 (2005).
5. P. M. Shearer, *Nature* **344**, 121 (1990).
6. J. J. Ita, L. Stixrude, *J. Geophys. Res.* **97**, 6849 (1992).
7. T. Gasparik, *Contrib. Mineral. Petrol.* **124**, 139 (1996).
8. K. Hirose, Y. Fei, *Geochim. Cosmochim. Acta* **66**, 2099 (2002).
9. K. D. Litasov, E. Ohtani, *Phys. Earth Planet. Inter.* **150**, 239 (2005).
10. Y. Nishihara, E. Takahashi, *Earth Planet. Sci. Lett.* **190**, 65 (2001).
11. G. R. Helffrich, B. J. Wood, *Nature* **412**, 501 (2001).
12. D. J. Weidner, Y. Wang, in *Earth's Deep Interior: Mineral Physics and Tomography From the Atomic to the Global Scale* [Geophys. Monogr. 117 Am. Geophys. Union (2000)], pp. 215–235.
13. Materials and methods are available on Science Online.
14. D. J. Frost, *Earth Planet. Sci. Lett.* **216**, 313 (2003).
15. K. Matsuzaka, M. Akaogi, T. Suzuki, T. Suda, *Phys. Chem. Miner.* **27**, 310 (2000).
16. At equilibrium the standard state free energy change $\Delta G_{p,T}^0$ for the reaction in Eq. 1 is related to the activities of the reacting components by the equation $\Delta G_{p,T}^0 = -RT \ln K = -RT \ln \left[\frac{a_{\text{Mg}_2\text{Al}_2\text{Si}_2\text{O}_7}^{\text{Gt}}}{\left[a_{\text{Ca}_2\text{Al}_2\text{Si}_2\text{O}_7}^{\text{Gt}} \right] \left[a_{\text{Mg}_2\text{Si}_2\text{O}_7}^{\text{Gt}} \right]^2} \right]$, where R is the gas constant, K is the equilibrium constant, and the superscript Gt refers to garnet. The component activities are described by $a_{\text{Mg}_2\text{Si}_2\text{O}_7}^{\text{Gt}} = \gamma_{\text{Mg}_2\text{Si}_2\text{O}_7}^{\text{Gt}} (X_{\text{Maj}}^{\text{Gt}})^3$, $a_{\text{Ca}_2\text{Al}_2\text{Si}_2\text{O}_7}^{\text{Gt}} = \gamma_{\text{Ca}_2\text{Al}_2\text{Si}_2\text{O}_7}^{\text{Gt}} (1 - X_{\text{Maj}}^{\text{Gt}}) (X_{\text{Ca}}^{\text{dodec}})^3$ and $a_{\text{Mg}_2\text{Al}_2\text{Si}_2\text{O}_7}^{\text{Gt}} = \gamma_{\text{Mg}_2\text{Al}_2\text{Si}_2\text{O}_7}^{\text{Gt}} (1 - X_{\text{Maj}}^{\text{Gt}}) (X_{\text{Mg}}^{\text{dodec}})^3$, where γ is an activity coefficient, x is the component mole fraction, and oct and dodec refer to the octahedral and dodecahedral cation sites of garnet and majorite (Maj). After calculating activity coefficients for mixing on just the dodecahedral site with use of a symmetric ternary solution model (for Mg, Ca, and Fe), we fit calculated values for the $RT \ln K$ term to obtain the equation $\Delta G_{p,T}^0 = 140763 + 26.7737T - 12560P$. The ternary solution model uses the Margules interaction parameters $W_{\text{FeMg}} = 300$ (J/mol) and $W_{\text{CaFe}} = 2000$ (J/mol) taken from the literature (17), whereas the parameter $W_{\text{MgCa}} = 8000 + 300P$ (J/mol, P in GPa) is refined by using the experimental data (13). Trial refinements showed no improvement to the fit of the data when nonideal mixing on the garnet octahedral site was included or when a reciprocal solution model was used.
17. H. St. C. O'Neill, B. J. Wood, *Contrib. Mineral. Petrol.* **70**, 59 (1979).
18. W. F. McDonough, S.-s. Sun, *Chem. Geol.* **120**, 223 (1995).
19. G. R. Helffrich, B. J. Wood, *Geophys. J. Int.* **126**, F7 (1996).
20. L. Stixrude, *J. Geophys. Res.* **102**, 14835 (1997).
21. L. Li et al., *Phys. Earth Planet. Inter.* **155**, 249 (2006).
22. S. V. Sinogeikin, J. D. Bass, *Geophys. Res. Lett.* **28**, 4335 (2001).
23. S. V. Sinogeikin, J. D. Bass, *Geophys. Res. Lett.* **29**, 10.1029/2001GL013937 (2002).
24. A. Suzuki et al., *Geophys. Res. Lett.* **27**, 803 (2000).
25. A. M. Dziewonski, D. L. Anderson, *Phys. Earth Planet. Inter.* **25**, 297 (1981).
26. P. M. Shearer, *J. Geophys. Res.* **101**, 3053 (1996).
27. M. P. Flanagan, P. M. Shearer, *J. Geophys. Res.* **103**, 2673 (1998).
28. U. R. Christensen, *J. Geophys. Res.* **102**, 22435 (1997).
29. H. Káráson, R. D. Van der Hilst, in *The History and Dynamics of Global Plate Motion* (Geophys. Monogr. 121 Am. Geophys. Union (2000)), pp. 277–288.
30. We thank D. Dolejš and S. Jacobsen for numerous useful discussions. A.S. was funded through the Elitenetzwerk Bavaria, International Graduate School on "Structure, Reactivity and Properties of Oxide Materials."

Supporting Online Material

www.sciencemag.org/cgi/content/full/319/5869/1515/DC1
Materials and Methods
Tables S1 to S4

9 November 2007; accepted 24 January 2008
10.1126/science.1152818

The Chlorine Isotope Composition of Earth's Mantle

M. Bonifacie,^{1,2} N. Jendrzewski,¹ P. Agrinier,¹ E. Humler,^{3,4} M. Coleman,^{5,6} M. Javoy¹

Chlorine stable isotope compositions ($\delta^{37}\text{Cl}$) of 22 mid-ocean ridge basalts (MORBs) correlate with Cl content. The high- $\delta^{37}\text{Cl}$, Cl-rich basalts are highly contaminated by Cl-rich materials (seawater, brines, or altered rocks). The low- $\delta^{37}\text{Cl}$, Cl-poor basalts approach the composition of uncontaminated, mantle-derived magmas. Thus, most or all oceanic lavas are contaminated to some extent during their emplacement. MORB-source mantle has $\delta^{37}\text{Cl} \leq -1.6$ per mil (‰), which is significantly lower than that of surface reservoirs (~ 0 ‰). This isotopic difference between the surface and deep Earth results from net Cl isotopic fractionation (associated with removal of Cl from the mantle and its return by subduction over Earth history) and/or the addition (to external reservoirs) of a late volatile supply that is ^{37}Cl -enriched.

Chloride is the major anion in most geological fluids. Chlorine is volatile, incompatible during silicate melting, and water soluble. For these reasons, geological processes—including partial melting, magma degassing, hydrothermal activity, and weathering—have concentrated Cl at the surface, particularly in the ocean, evaporites, and crustal brines. Considering only the MORB-source mantle and surface reservoirs, we can estimate that the former, with 1 to 8 parts per million (ppm) of Cl (1–3), contributes a maximum of $\sim 20\%$ to the Cl budget and perhaps much less, given the difficulty in obtaining a complete inventory of crustal brines. Earth's Cl budget may thus differ from that of most other abundant and geochemically important volatiles (e.g., H_2O , CO_2).

There are several first-order questions about Cl in Earth: What early solar system processes determined Earth's initial Cl budget? How, and when, was most of Earth's Cl extracted from the mantle? How much Cl is returned to the mantle by subduction? Is the extraction of Cl from, and its return to, the mantle in steady state? If not, what can we learn about the history of Earth's volatile cycles by reconstructing secular changes in the Cl budgets of the mantle and/or surface? These questions can be addressed by determining $\delta^{37}\text{Cl}$ (4) of the various reservoirs in that budget and isotopic fractionations associated with exchanges between those reservoirs. Surface reservoirs (mostly oceans, brines, and evaporites) are known to have a collective mean $\delta^{37}\text{Cl}$ of 0.0 ± 0.5 ‰ (5–8). However, it is more difficult to

measure the $\delta^{37}\text{Cl}$ of pristine mantle rocks and mantle-derived magmas because both are generally poor in Cl (1–3), which is usually present in insoluble form. Here, we present measurements of 22 MORBs that define the mantle composition to lower $\delta^{37}\text{Cl}$ values than previously assumed (9, 10) and use these results to examine the origin and history of Earth's volatiles.

The first attempt to define the Cl isotope composition of the mantle used thermal ionization mass spectrometry (TIMS) to analyze three MORB samples (9), yielding an average mantle value of $+4.7$ ‰. Sharp et al. (10) recently used gas-source mass spectrometry to determine a mantle $\delta^{37}\text{Cl}$ value of -0.1 ± 0.4 ‰, which was derived from the average of 11 MORB samples, two kimberlites, and three carbonatites. They suggested $\delta^{37}\text{Cl}$ homogeneity near the seawater value (0‰) for mantle and surface reservoirs over Earth's history (10). This implies that neither removal of Cl from the mantle nor its

¹Equipe de Physico-Chimie des Fluides, Géologiques, Institut de Physique du Globe de Paris (IPGP) and Université Paris, Diderot (UPD), CNRS-INSU—Unité Mixte de Recherche (UMR) 7154, 75251 Paris Cedex 05, France. ²Division of Geological and Planetary Sciences, California Institute of Technology (CalTech), Pasadena, CA 91125, USA. ³Equipe de Géosciences Marines, IPGP, CNRS-INSU-UMR 7154, 75251 Paris Cedex 05, France. ⁴Laboratoire de Planétologie et Géodynamique, Université de Nantes, CNRS UMR 6112, 44322 Nantes Cedex 3, France. ⁵School of Human and Environmental Sciences, University of Reading, RG6 6AB, UK. ⁶Jet Propulsion Laboratory, CalTech, Pasadena, CA 91109, USA.

*To whom correspondence should be addressed. E-mail: bonifaci@gps.caltech.edu

return by subduction led to substantial net Cl isotopic fractionation (isotopic steady state). Such a conclusion is consistent with the apparent (i) lack of secular change in $\delta^{37}\text{Cl}$ in sediments (10) [although data include large, unexplained variations about the long-term mean, from -3.2 to $+2.5\%$ (10)] and (ii) lack of Cl isotope fractionation through the subduction-related partial loss of Cl (11) inherited from oceanic alteration. However, Cl isotope fractionation of up to 8% has been seen in surface reservoirs during alteration of the oceanic lithosphere (11–15) and/or in sedimentary settings (16–18).

We determined the Cl contents and isotopic compositions of 22 MORB samples using a low-blank method (15, 19) involving the extraction of structurally bound Cl by pyrohydrolysis, followed by gas-source, dual-inlet mass spectrometric measurements. Analytical yields average $100 \pm 8\%$, blanks are typically 7% of the sample chloride analyzed, and the whole-procedure external precision for $\delta^{37}\text{Cl}$ measurements (1σ) is $\pm 0.14\%$ (fig. S1). Our samples are nominally zero-age lavas collected from the Pacific, Atlantic, and Indian oceans and cover large ranges in extent of fractional crystallization [MgO, 5.9 to 9.2 weight percent (wt %)], partial melting [Na_2O corrected for crystal fractionation to 8% MgO ($\text{Na}_{8,0}$), 1.9 to 4.2 wt %], source enrichment (K/Ti, 0.04 to 0.41), eruption depths [1800 to 4905 meters below sea level (mbsl)], spreading rate (16 to 153 mm yr^{-1}), and degree of Cl contamination (K/Cl, 1.0 to 38.3) (table S1). There is substantial evidence that some terrestrial lavas are contaminated with surficial Cl to some extent (1, 2, 20, 21) because of the enormous contrast in Cl concentrations between primary magmas (about <80 ppm) and surficial waters and rocks (typically 1000s ppm). Our sample suite spans a relatively large range in Cl concentration (42 to 701 ppm), so we can examine how the samples vary with isotopic composition.

Our $\delta^{37}\text{Cl}$ values for MORB glasses range from -1.9 to -0.1% and increase with Cl concentration (Fig. 1). The most Cl-rich samples have $\delta^{37}\text{Cl}$ values close to seawater [$\delta^{37}\text{Cl} = 0\%$ (5, 6)], whereas Cl-poor MORBs have substantially lower $\delta^{37}\text{Cl}$ values.

One hypothetical explanation of the linear $\delta^{37}\text{Cl}$ versus $1/\text{Cl}$ ($\delta^{37}\text{Cl}$ - $1/\text{Cl}$) trend is that all MORBs have $\delta^{37}\text{Cl}$ values near 0% , and our analytical technique contributed a low- $\delta^{37}\text{Cl}$ blank that decreased the measured values of Cl-poor samples. Our analytical blank has a $\delta^{37}\text{Cl}$ value of $-0.3 \pm 2.2\%$ (fig. S2), which is far above the value of $-9.8 \pm 4.3\%$ needed to explain the trend in Fig. 1. Thus, the $\delta^{37}\text{Cl}$ - $1/\text{Cl}$ trend reflects real, coupled variations in Cl concentration and isotopic composition of MORBs, not laboratory contamination. Because Cl is highly incompatible [partition coefficient $D_{\text{therzolite/silicate-melt}} \sim 0.002$ (22)] and does not undergo speciation changes by partitioning between silicate minerals and melts during crystallization or melting, variable degrees of partial melting or extents

of crystallization should produce only negligible Cl isotope fractionation at magmatic temperatures (23, 24). Because our MORBs were collected from relatively great depths (≥ 1800 mbsl) and have relatively low Cl contents (≤ 701 ppm), they remained undersaturated with Cl during eruption (25, 26) and should not have lost Cl by degassing. Thus, the $\delta^{37}\text{Cl}$ - $1/\text{Cl}$ trend reflects

mixing between two (or more) isotopically distinct components that are commonly sampled during the generation, differentiation, and/or eruption of MORBs.

The $\delta^{37}\text{Cl}$ - $1/\text{Cl}$ trend might reflect mixing between primary mantle Cl and Cl added pre- or syn-eruptively to those magmas, such as what might result from contamination (assimilation) of mag-

Fig. 1. $\delta^{37}\text{Cl}$ variation with the reciprocal Cl content $\times 1000$ (ppm^{-1}) for MORB glasses determined by gas-source mass spectrometry. Large closed circles indicate nominally submarine on-axis MORB from the present study: N-MORB, black circles; E-MORB, gray circles. Small open symbols represent Sharp *et al.* data (10): nominally submarine on-axis MORB, circles; submarine off-axis samples, triangles; aerial sample, square. Seawater and ridge-axis hydrothermal brines are denoted by the black star (5, 6, 27), and the $\delta^{37}\text{Cl}$ range of altered oceanic rocks is denoted by the vertical gray arrow (11). For our study, the weighted least-squares fit for all MORB ($y = -0.052x - 0.378$; $R^2 = 0.72$) is shown (dashed line). Error bars on Cl contents (1σ) are indicated when larger than the symbol. Error bars on $\delta^{37}\text{Cl}$ values (1σ) represent the whole-procedure external uncertainty $\pm 0.14\%$ (determined on repeat extractions and analyses of various amounts of our internal reference-rock standard SO100DS92) (19). For the Sharp *et al.* data (10): reported Cl contents are those acquired by independent measurements of glass fragments by ion or electron microprobe; no uncertainties on Cl content were reported; error bars on $\delta^{37}\text{Cl}$ data (one analysis per sample) represent the $\delta^{37}\text{Cl}$ -measurement errors determined on replicate analyses of seawater samples, either by continuous-flow or dual-inlet modes (± 0.26 and $\pm 0.10\%$, respectively), not total analytical uncertainty for silicate samples with structurally bound Cl.

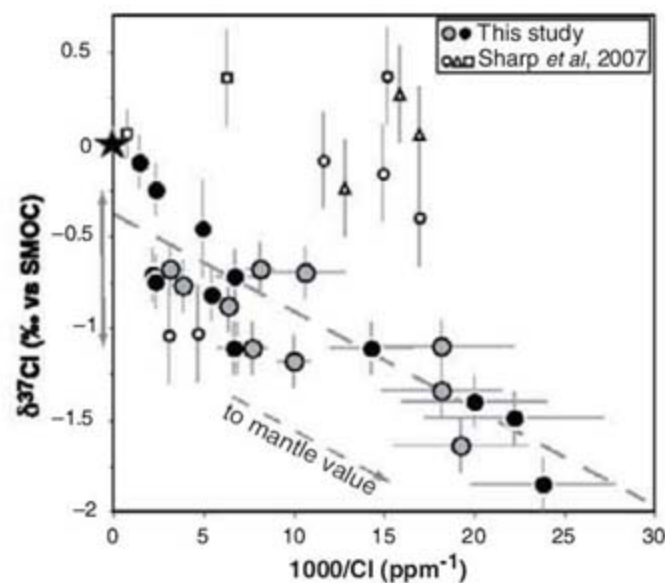
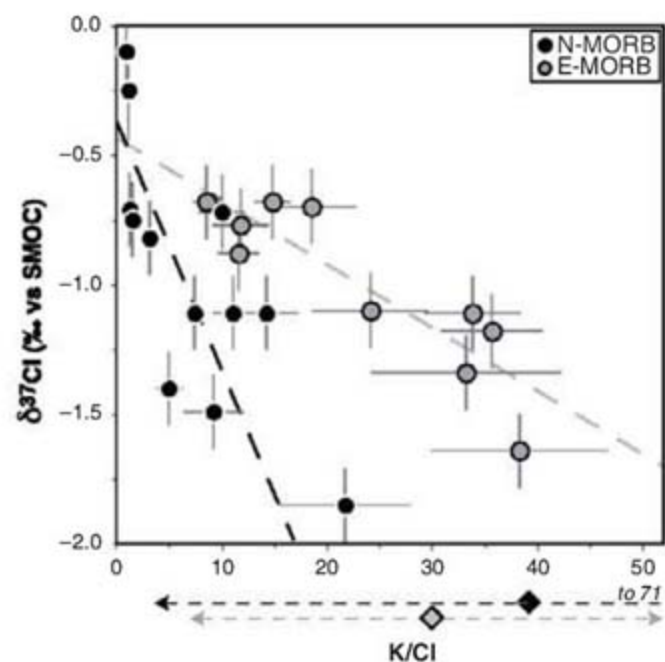


Fig. 2. $\delta^{37}\text{Cl}$ variation with K/Cl ratios for MORB glasses analyzed in this study. N-MORB, black circles; E-MORB, gray circles. The weighted least-squares regression fits for N-MORB [$y = -0.097x - 0.354$ ($R^2 = 0.58$); black dashed line] and E-MORB [$y = -0.024x - 0.430$ ($R^2 = 0.78$); gray dashed line] are shown. Error bars on K/Cl (1σ) are indicated when larger than the symbol. K/Cl average values (diamonds) and the 1σ intervals (dashed arrows) calculated from data on N-MORBs and E-MORBs (black and gray, respectively) documented in (2) and expected to not have experienced contamination [because they were collected in areas with low magma flux (i.e., slow spreading rate and low extent of melting)]. The consistency of the trends observed here (as well as in Fig. 1) from the least (lowest $\delta^{37}\text{Cl}$) to the most contaminated samples for both suites—and even for samples showing $\text{K/Cl} \geq 12.5$, a value previously suggested to reflect the lower limit for uncontaminated MORB (2)—suggests that it is perhaps inevitable that the Cl budget of MORB is influenced by such contamination. Accordingly, $\delta^{37}\text{Cl}$ values of MORB would be a more sensitive index of the extent of contamination than K/Cl values.



mas by a Cl-rich, seawater-derived material (e.g., seawater, brines, and altered rocks) (1, 2, 20, 21). According to this hypothesis, the Cl-poor, low- $\delta^{37}\text{Cl}$ end of the trend (Fig. 1) most closely approaches the composition of primary MORB, and increasing degrees of mixing with seawater-derived Cl increase their Cl contents and draw the $\delta^{37}\text{Cl}$ values of the most contaminated lavas toward the value of seawater. Several independent lines of evidence support such an interpretation: (i) the high- $\delta^{37}\text{Cl}$ end-member of the mixing trend is within the range of compositions of seawater or on-axis high-temperature hydrothermal fluids [both $\sim 0\%$ (5, 6, 27)] and altered oceanic rocks [about $-0.7 \pm 0.4\%$ (11)]; (ii) the Cl contents of our lowest $\delta^{37}\text{Cl}$ samples imply a typical mantle Cl concentration of less than 4 ± 3 ppm [assuming MORBs are typically produced by 10% mantle melting and Cl is highly incompatible during mantle melting (22)], within the previously suggested range (1 to 8 ppm) (1–3); and (iii) MORBs from relatively fast-spreading ridges, expected to have experienced larger extents of contamination (1, 2, 20, 21), show higher $\delta^{37}\text{Cl}$ values (closer to 0‰) than MORBs from relatively slow-spreading ridges (averages for Pacific and Indian samples of -0.8 and -1.4% , respectively). Because K is an incompatible element and is enriched in lavas due to assimilation far less than Cl is, K/Cl ratios have been used as an index of assimilation (1, 2, 20). Thus, the fact that our samples with the lowest K/Cl values also have the highest $\delta^{37}\text{Cl}$ (and vice versa; samples SO22-17D and MD57D-7-2, respectively) supports our interpretations of the trend in Fig. 1 (i.e., high-Cl, high- $\delta^{37}\text{Cl}$ samples are more contaminated than those with low Cl and low $\delta^{37}\text{Cl}$). It is noteworthy that, whereas normal-MORBs (N-MORBs) and enriched-MORBs (E-MORBs) define a common trend in Fig. 1, they appear to define two distinct trends in a $\delta^{37}\text{Cl}$ versus K/Cl plot (Fig. 2): Starting at a common intercept at low K/Cl and high $\delta^{37}\text{Cl}$, K/Cl increases gradually with decreasing $\delta^{37}\text{Cl}$ in N-MORB and relatively strongly in E-MORB. Among the least contaminated samples (i.e., those with the lowest $\delta^{37}\text{Cl}$ values), E-MORBs have K/Cl ratios about twice those of N-MORBs at a given $\delta^{37}\text{Cl}$ value. It is possible that the distinction between the N- and E-MORB trends in Fig. 2 is fortuitous and would not be observed in a broader and more representative sampling. In particular, N- and E-MORBs have been distinguished here by their K/Ti ratios, though Ti is heterogeneous in the mantle (28, 29). Also, the distinct K/Cl ratios among our E- and N-MORB samples result mainly from higher K contents of E-MORBs, not from distinct Cl contents. Nevertheless, this result, taken at face value, could reflect differences in K/Cl ratio and/or $\delta^{37}\text{Cl}$ value between the mantle sources of E- and N-MORBs. At one extreme, the mantle sources of N- and E-MORBs may be comparable in $\delta^{37}\text{Cl}$, but E-MORBs are higher in K/Cl ratios than N-MORBs [contrary to what has been previously suggested; e.g., (2)]. One possible explanation for such a difference is that E-MORB sources

contain larger fractions of subducted material that are rich in K but relatively poor in Cl, perhaps because most of Cl in subducted slabs is returned to the surface by arc volcanism rather than being efficiently recycled. Alternatively, these two types of mantle sources might be similar in K/Cl ratio (~ 34 ; see diamonds in Fig. 2), but N-MORB sources are lower in $\delta^{37}\text{Cl}$ than E-MORB sources.

Assuming that our MORB samples with the lowest Cl concentrations are the least contaminated by seawater-derived Cl, we estimate that the MORB-source mantle (depleted and enriched parts) is characterized by a $\delta^{37}\text{Cl}$ value of -1.6% or less, which is lower than values previously proposed (9, 10). The different value of Magenheimer *et al.* (9) can be ascribed to analytical errors in the TIMS-based measurements (15). The cause of the difference from the results of Sharp *et al.* (10)—the methods of which are more similar (though not identical) to ours—is unclear. Assuming that both data sets are analytically accurate, it is difficult to ascribe the discrepancy to a difference in sample selection, because their suite includes some samples with relatively high $\delta^{37}\text{Cl}$ and low nominal Cl contents, which is incompatible with the $\delta^{37}\text{Cl}$ -I/Cl trend for our data (Fig. 1). It also seems unlikely that the diverse suite of nominally MORBs we examined (from the three main oceans and showing large variations in parameters such as Na_8O , K/Ti, K/Cl, and spreading rate) would yield the simple $\delta^{37}\text{Cl}$ -I/Cl trend that we found if low-Cl, high- $\delta^{37}\text{Cl}$ MORBs were common. Given that Sharp *et al.* (10) report neither Cl contents of basalts from the actual extractions of Cl for isotopic analysis nor analytical yields, a close comparison of the two data sets is not possible.

The coherence of the observed trends suggests that the mantle $\delta^{37}\text{Cl}$ value unmodified by contamination is $\leq -1.6\%$, which is significantly more negative than that of the surface reservoirs ($\sim 0\%$). This isotopic difference might result from (i) systematic ^{37}Cl enrichment in surface reservoirs and corresponding depletion in the mantle via long-term Cl extraction from and return to the mantle and/or (ii) a primary ^{37}Cl enrichment of the surface Cl of early Earth due to a late supply of volatile-rich material (30), which was enriched with ^{37}Cl as compared with the material from which Earth was mainly accreted. In the first case, because the mantle now contains at least one-fifth as much Cl as surface reservoirs, any $\delta^{37}\text{Cl}$ variations resulting from these long-term exchanges will be more important in the mantle or mantle-derived magmas than in the oceans or sedimentary rocks. The second hypothesis finds support from our current knowledge of the Cl cycle, because the subduction flux shows substantially higher $\delta^{37}\text{Cl}$ (preliminarily estimated from serpentinites, average $\delta^{37}\text{Cl} = -0.7 \pm 0.4\%$) (11) than the mantle flux ($\delta^{37}\text{Cl} \leq -1.6\%$); such a combination of fluxes could only have decreased the $\delta^{37}\text{Cl}$ gap between mantle and surface reservoirs over the long-term exchanges.

References and Notes

1. A. Jambon, B. Dérulle, G. Dreibus, F. Pineau, *Chem. Geol.* **126**, 101 (1995).
2. P. J. Michael, W. C. Cornell, *J. Geophys. Res.* **103**, 18325 (1998).
3. A. E. Saal, E. H. Hauri, C. H. Langmuir, M. R. Perfit, *Nature* **419**, 451 (2002).
4. Chlorine stable isotope compositions are expressed as $\delta^{37}\text{Cl} = \{[(^{37}\text{Cl}/^{35}\text{Cl})_{\text{sample}} / (^{37}\text{Cl}/^{35}\text{Cl})_{\text{SMOC}}] - 1\} \times 1000$, where SMOC is standard mean ocean chloride, defined as 0‰.
5. R. Kaufmann, A. Long, H. Bentley, S. Davis, *Nature* **309**, 338 (1984).
6. A. Godon *et al.*, *Chem. Geol.* **207**, 1 (2004).
7. H. G. M. Eggenkamp, R. Kreulen, A. F. Koster Van Groos, *Geochim. Cosmochim. Acta* **59**, 5169 (1995).
8. C. J. Eastoe, T. M. Peryt, O. Y. Petrychenko, D. Geisler-Cussey, *Appl. Geochem.* **22**, 575 (2007).
9. A. J. Magenheimer, A. J. Spivack, P. J. Michael, J. M. Gieskes, *Earth Planet. Sci. Lett.* **131**, 427 (1995).
10. Z. D. Sharp *et al.*, *Nature* **446**, 1062 (2007).
11. M. Bonifacie *et al.*, *Geochim. Cosmochim. Acta* **72**, 126 (2008).
12. Z. D. Sharp, J. D. Barnes, *Earth Planet. Sci. Lett.* **226**, 243 (2004).
13. M. Bonifacie *et al.*, *Geochim. Cosmochim. Acta* **69**, A787 (2005).
14. J. D. Barnes, Z. D. Sharp, *Chem. Geol.* **228**, 246 (2006).
15. M. Bonifacie *et al.*, *Chem. Geol.* **242**, 187 (2007).
16. A. Godon *et al.*, *Geochim. Cosmochim. Acta* **68**, 2153 (2004).
17. M. Bonifacie, C. Monnin, N. Jendrzewski, P. Agrinier, M. Javoy, *Earth Planet. Sci. Lett.* **260**, 10 (2007).
18. V. Woulé Ebongué, N. Jendrzewski, F. Walgenwitz, F. Pineau, M. Javoy, *AAPG Bull.* **89**, 1005 (2005).
19. Materials and methods are available as supporting material on Science Online.
20. P. J. Michael, J.-G. Schilling, *Geochim. Cosmochim. Acta* **53**, 3131 (1989).
21. S. A. Soule *et al.*, *Earth Planet. Sci. Lett.* **252**, 289 (2006).
22. J.-G. Schilling, M. B. Bergeron, R. Evans, *Philos. Trans. R. Soc. London Ser. A* **297**, 147 (1980).
23. P. Richey, Y. Bottinga, M. Javoy, *Annu. Rev. Earth Planet. Sci.* **5**, 65 (1977).
24. H. P. Schauble, G. R. Rossman, H. P. Taylor Jr., *Geochim. Cosmochim. Acta* **67**, 3267 (2003).
25. C. K. Unni, J.-G. Schilling, *Nature* **272**, 19 (1978).
26. J. D. Webster, R. J. Kinzler, E. A. Mathez, *Geochim. Cosmochim. Acta* **63**, 729 (1999).
27. M. Bonifacie, J. L. Charlou, N. Jendrzewski, P. Agrinier, J. P. Donval, *Chem. Geol.* **221**, 279 (2005).
28. C. H. Langmuir, E. M. Klein, T. Plank, in *Mantle Flow and Melt Generation at Mid-Ocean Ridges*, J. Phipps-Morgan, D. K. Blackman, J. M. Sinton, Eds. (American Geophysical Union, Washington, DC, 1992), vol. 71 of Geophysical Monograph Series, pp. 183–280.
29. C. M. Meyzen, M. J. Toplis, E. Humler, J. N. Ludden, C. Mével, *Nature* **421**, 731 (2003).
30. M. Javoy, *Chem. Geol.* **147**, 11 (1998).
31. We thank R. Hékinian, A. Jambon, C. Mével, M. Moreira, and F. Pineau for providing most of the samples, as well as A. Bézos, P. Cartigny, J. M. Eiler, F. Guyot, C. H. Langmuir, and M. Moreira for discussion on this work and, together with other reviewers, for their comments on a previous version of this manuscript. M.B. thanks J. M. Eiler and E. M. Stolper at CalTech for their support during the writing of this manuscript. This is IGP contribution 2342 and CNRS-INSU contribution 408.

Supporting Online Material

www.sciencemag.org/cgi/content/full/319/5869/1518/DC1

Materials and Methods

Figs. S1 and S2

Tables S1 and S2

References

26 September 2007; accepted 1 February 2008
10.1126/science.1150988

Caribbean Reef Development Was Independent of Coral Diversity over 28 Million Years

Kenneth G. Johnson,^{1*} Jeremy B. C. Jackson,^{2,3} Ann F. Budd⁴

The relationship between natural variations in coral species diversity, reef development, and ecosystem function on coral reefs is poorly understood. Recent coral diversity varies 10-fold among geographic regions, but rates of reef growth are broadly similar, suggesting that diversity is unimportant for reef development. Differences in diversity may reflect regional differences in long-term biotic history in addition to environmental conditions. Using a combination of new and published fossil and stratigraphic data, we compared changes in coral diversity and reef development within the tropical western Atlantic over the past 28 million years. Reef development was unrelated to coral diversity, and the largest reef tracts formed after extinction had reduced diversity by 50%. High diversity is thus not essential for the growth and persistence of coral reefs.

Coral reefs are among the most diverse and productive ecosystems in modern oceans (1, 2). Scleractinian corals form the primary wave-resistant structural reef framework that provides complex three-dimensional habitat for diverse associated biota, and through their skeletal growth, they are responsible for reef development that typically results in the accumulation of massive sequences of carbonate rock over tens of thousands of years (1). The recent decline of coral abundance because of human impacts (3, 4) has focused attention on the potential loss of biodiversity (4) and the assumed importance of high coral diversity for coral reef persistence and resilience (5, 6). However, reef development is controlled globally primarily by environmental conditions, with the most extensive reefs being restricted to tropical or subtropical regions with clear, warm surface water that is supersaturated with respect to aragonite (7). In contrast, variations in species diversity reflect long-term regional differences in patterns of species origination and extinction as well as environmental conditions (8).

Coral diversity on Indo-Pacific reefs is 10 times higher than on Caribbean reefs (9), but the rates of carbonate production and reef growth are similar (10), suggesting that coral diversity is unimportant to reef development. High-latitude reefs with low coral diversity, such as at Lord Howe Island in the Pacific (11), Clipperton Atoll in the eastern Pacific (12), and Bermuda in the Atlantic (13), also exhibit strong reef development. Hence, we must ask whether diversity affects reef development within a single biogeographic region independent of confounding

differences in regional biotic history and latitude. To this end, we examined the relationship between reef coral diversity and coral reef development in the tropical western Atlantic during the past 28 million years.

The historical waxing and waning of reef development over tens of millions of years throughout the tropical western Atlantic is well established (14, 15), but the distribution of fossil reef coral species was inadequately known until recently (16, 17). We used published descriptions of the stratigraphic distribution and relative thickness of diverse facies of carbonate units associated with coral frameworks as a first-order proxy for reef development (18). Collections were grouped into 46 faunules that comprise sets of collections from the same stratigraphic horizon and locality (table S3). The stratigraphic coverage of faunules is complete within intervals from the Late Oligocene to Recent (fig. S1). Specimens were identified as to species on the basis of quantitative metrics of coral skeletal morphology that are consistent with new molecular systematics. These data were supplemented by published occurrences from six widely distributed Late Pleistocene deposits. The resulting compilation includes 17,306 identified specimens from 3960 occurrences and 885 collections or published lists (table S2). Unlike other literature-based compilations (19), species occurrences are based mainly on specimens systematically collected from old and new exposures within a revised stratigraphic and paleoenvironmental framework with a temporal precision of just a few hundred thousand years in most cases (18). This comprehensive coverage within a 28-million-year window in a single geographic region allows more precise estimates of the species richness of reef corals within each sample unit.

Fossil data on reef development and species occurrences were divided into seven stratigraphic intervals (Fig. 1) (18). The seven geological intervals generally correspond to sub-epoch boundaries, with the exception of the Late Pliocene through Early Pleistocene interval that extends

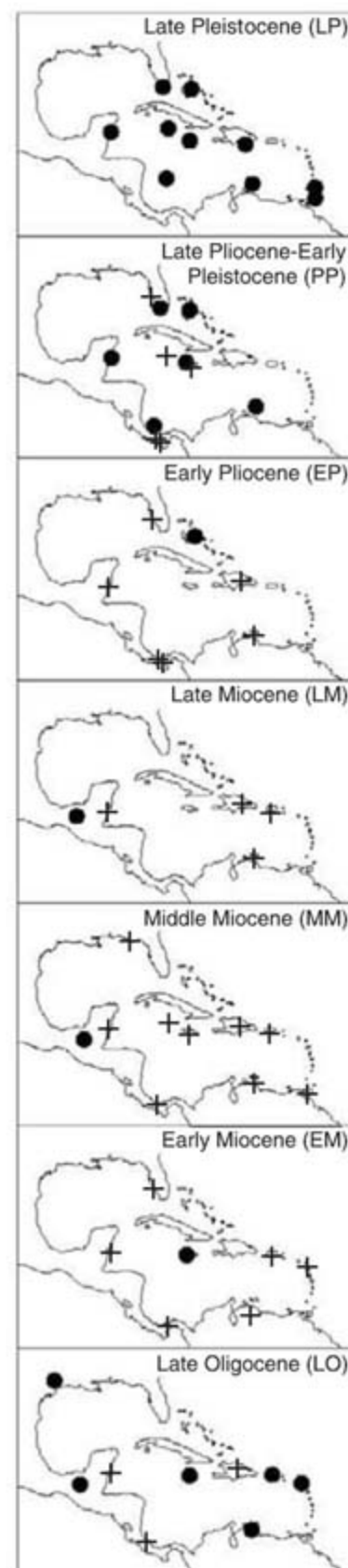


Fig. 1. Distribution of Caribbean fossil coral reefs through time from the Late Oligocene to Late Pleistocene. Stratigraphic units including extensive reef buildups are indicated by = symbols, and units containing abundant zooxanthellate corals but without extensive reef building are indicated by + symbols. The lower boundaries for stratigraphic bins are defined as Late Oligocene = 28.45 Ma, Early Miocene = 23.03 Ma, Late Miocene = 11.61 Ma, Early Pliocene = 5.33 Ma, Late Pliocene–Early Pleistocene = 1.6 Ma, and Late Pleistocene = 1 Ma (18).

¹Department of Palaeontology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK. ²Center for Marine Biodiversity and Conservation, Scripps Institution of Oceanography, La Jolla, CA 92093, USA. ³Smithsonian Tropical Research Institute, Balboa, Republic of Panama. ⁴Department of Geoscience, University of Iowa, Iowa City, IA 52242, USA.

*To whom correspondence should be addressed. E-mail: K.Johnson@nhm.ac.uk

from 3.6 to 1.0 million years ago (Ma) (18). There were two intervals of extensive reef development within the tropical western Atlantic over the past 28 million years (Fig. 1). The first, in the Late Oligocene, is represented by thick carbonate platforms in Antigua, Jamaica, Puerto Rico, Chiapas, Venezuela, and the Gulf of Mexico, and the second occurs in the Late Pliocene and Pleistocene, when extensive reef tracts developed throughout the entire region, especially over the past 2 million years. During the intervening 20 million years, carbonate accumulation was more highly localized and corals occurred in small patches in siliciclastic settings characteristic of the northern Dominican Republic, the Anguilla Formation, and the Bocas del Toro and Limon Basins of lower Central America. Moreover, thick Miocene carbonates, such as the Tamara Formation in Trinidad, have locally abundant corals but no evidence of extensive coral framework (table S5).

The shapes of the collecting curves for reef coral species within each of the seven stratigraphic intervals suggest that few new species remain to be discovered in rocks from most of the intervals (Fig. 2A). Diversity was highest from the Late Miocene to Late Plio-Pleistocene and peaked in the Early Pliocene. Diversity was lowest during the Late Oligocene to early Miocene and Late Pleistocene, and there was strong agreement between observed and estimated diversity except for the high value of the Chao 2 estimator within the Early Miocene. Confidence intervals on diversity (Fig. 2B) were estimated using the O2W subsampling method (20). There is no correlation between coral diversity and the geographic extent of reef development over the seven intervals (Fig. 3) (Pearson's $R = -0.325$, $df = 5$, $P = 0.641$). Most strikingly, the Late Pleistocene interval that exhibits the most extensive Caribbean coral reef development in the past 28 million years is characterized by exceptionally low species diversity of reef-building corals. The numbers of coral species plummeted sometime toward the end of the Late Pliocene to Early Pleistocene interval, when half of the Late Pli-

ocene species became extinct (21). Yet, in spite of this massive extinction, reef development dramatically increased.

Ordination of faunules with 10 or more species was carried out using both principal-coordinates analysis (PCA) and nonmetric multidimensional scaling for both presence-absence and abundance data (18). Results were very similar and are presented only for PCA based on presence-absence data (Fig. 4). The first two PCA axes explain 58 and 10% of the variance, so that the two-dimensional representation in Fig. 3 explains two-thirds of the total variance in community composition. Unsurprisingly, the largest breaks in community composition are associated with the periods of strongest faunal turnover and changes in reef development. There are distinct gaps in community composition between the Late Oligocene and Early Miocene and between the Early and Middle Miocene, suggesting major ecological shifts associated with the Early Miocene extinction that reduced coral species diversity by about 40% and coincided with a dramatic decrease in reef development (17, 22) (Figs. 1 and 2B). There is also strong separation between the Middle Miocene through Early Pliocene intervals, when diversity was rising but reef development was minimal, and the Late Pliocene through Pleistocene, when reef development dramatically increased despite mass extinction of coral species.

The independence of reef development from coral species diversity (Fig. 3) and the striking increase in Pleistocene reef development (Fig. 1) closely following the massive extinction of the Late Pliocene coral fauna (Fig. 2) strongly suggest that factors other than coral diversity are more important to reef building. Thus, our results support the implication of Recent biogeographic distributions of corals and coral reefs that high coral diversity is unnecessary for vigorous reef growth. Another likely example from the fossil record includes the thick Miocene reef deposits in the Mediterranean that were built by fewer than 10 species of corals (23, 24). In contrast, Kiessling observed strong positive correlations between

reef diversity and the stability of reef structure, architecture, and construction style over the entire Phanerozoic (19). However, the great differences in scale, stratigraphic and taxonomic resolution, and parameters measured between our studies complicate meaningful comparison.

Extensive reef development in the modern ocean is controlled mainly by temperature and surface productivity (25). The physical environment sets the boundary conditions for the biota and determines accumulation rates of calcium carbonate and ecological zonation, because the coral symbiosis functions most effectively within a limited range of nutrient and temperature conditions (26). Thus, under the appropriate environmental conditions, the production and accumulation of carbonate by both low- and high-diversity coral communities can result in substantial reef development. This conclusion is further supported by the opposite consequences for reef development after the two episodes of high coral extinction in the Early Miocene and Pleistocene, both of which are associated with strong regional changes in surface ocean productivity (27, 28). Paleooceanographic modeling and the selectivity of changes in the reef coral fauna imply that regional planktonic productivity strongly increased from the Late Oligocene to Middle Miocene, when extensive reef development collapsed (29). In contrast, a variety of data, including a 50% drop in the mean annual range in temperature (30), increased percent of carbonate in sediments (30), and changes in the taxonomic composition of common benthic foraminifera (31) all point toward a strong regional decline in surface productivity associated with the increasing isolation of the Caribbean Basin from eastern Pacific bottom waters during the gradual uplift of the Isthmus of Panama (32).

Low surface productivity allowed the development of extensive Caribbean reef tracts in the Late Oligocene and Quaternary, but the structural composition of reef communities was very different in these two intervals. Late Oligocene coral framework is dominated by heavily calcified and massive coral colonies (8), whereas Quaternary communities are dominated by more

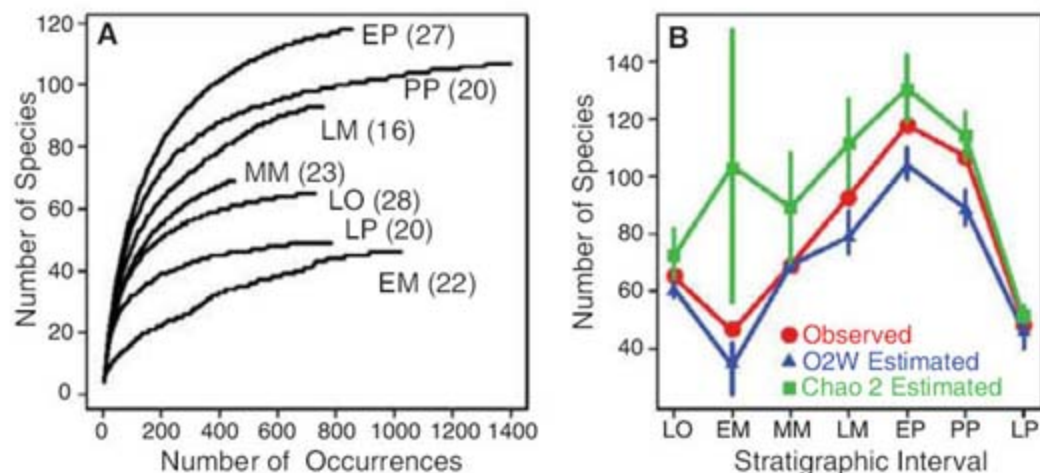


Fig. 2. Diversity of reef coral species for seven intervals from the Late Oligocene to Late Pleistocene. (A) Cumulative collection curves for each interval (numbers of faunules per interval are indicated in parentheses). (B) Observed, resampled (O2W-estimated), and Chao 2-estimated measures of species diversity per interval over geological time.

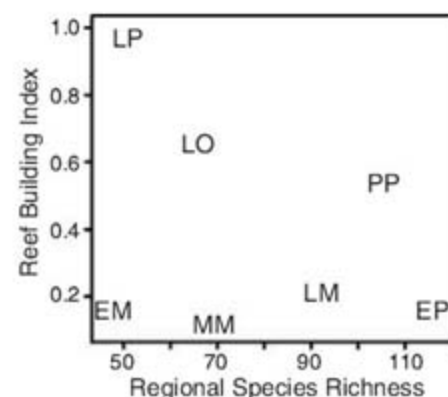


Fig. 3. Resampled coral species diversity versus the proportional regional distribution of extensive coral reef development for the seven sample intervals.

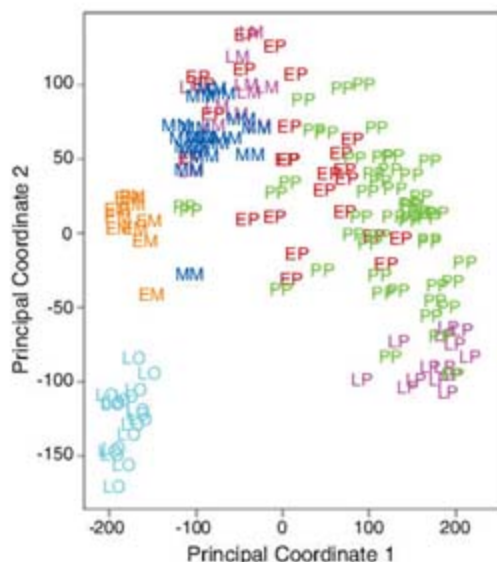


Fig. 4. PCA of coral community composition for the 46 faunules identified as to membership in the different sample intervals.

porous, frond-shaped *Acropora* species that grow an order of magnitude faster than any other corals (33). *Acropora* gradually increased in abundance during the Pliocene as slower-growing *Stylophora* and *Pocillopora* species became extinct (22, 34, 35). The rise of *Acropora* is correlated with, and perhaps caused by, great differences in rates of sea-level fluctuations between the Late Oligocene and Quaternary (36). Exceptionally rapid growth and high rates of fragmentation by *Acropora* species allowed them to keep up with a sudden rise in sea level, as has occurred repeatedly during the past 1 million years (37). Acroporids were present on Oligocene reefs (8, 14, 17) but failed to achieve ecological dominance until the Pleistocene (38).

Thus, the ecological characteristics of dominant coral species are more important to reef growth and development than the simple number of species. This is all too apparent in the precipitous decline of Caribbean reef communities and reef development in relation to the extreme Caribbean-wide decline in *Acropora* abundance due to human disturbance (33, 39).

References and Notes

- B. G. Hatcher, *Trends Ecol. Evol.* **5**, 149 (1990).
- N. Knowlton, *Proc. Acad. Nat. Sci. U.S.A.* **98**, 5419 (2001).
- T. A. Gardner, I. M. Cote, J. A. Gill, A. Grant, A. R. Watkinson, *Science* **301**, 958 (2003).
- J. F. Bruno, E. R. Selig, *PLoS ONE* **2**, e711 (2007).
- D. R. Bellwood, T. P. Hughes, C. Folke, M. Nyström, *Nature* **429**, 827 (2004).
- D. R. Bellwood, T. P. Hughes, *Science* **292**, 1532 (2001).
- J. A. Kleypas, J. W. McManus, L. A. B. Meñez, *Am. Zool.* **39**, 146 (1999).
- A. F. Budd, *Coral Reefs* **19**, 25 (2000).
- J. E. N. Veron, *Corals of the World* (Australian Institute of Marine Science, Townsville, Australia, 2000).
- W.-C. Dullo, *Facies* **51**, 37 (2005).
- V. J. Harriott, S. A. Banks, *Coral Reefs* **21**, 83 (2002).
- P. W. Glynn, J. E. N. Veron, G. M. Wellington, *Coral Reefs* **15**, 71 (1996).
- A. Logan, *Sedimenta* **XI**, 1 (1988).
- T. W. Vaughan, *U.S. Natl. Mus. Bull.* **103**, 189 (1919).

- C. Perrin, in *Phanerozoic Reef Patterns*, W. Kiessling, E. Flügel, J. Golonka, Eds. [Society for Sedimentary Geology (SEPM) Special Publication 72, SEPM, Tulsa, OK, 2002], pp. 587–621.
- K. G. Johnson, in *Fossil Corals and Sponges: Proceedings of the Ninth International Symposium on Fossil Cnidaria and Porifera*, B. Hubmann, W. E. Piller, Eds. (Austrian Academy of Sciences, Vienna, Austria, 2007), pp. 471–491.
- A. F. Budd, T. A. Stemann, K. G. Johnson, *J. Paleontol.* **68**, 951 (1994).
- Materials and methods are available as supporting material on Science Online.
- W. Kiessling, *Nature* **433**, 410 (2005).
- J. Alroy et al., *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6261 (2001).
- A. F. Budd, K. G. Johnson, *Paleobiology* **25**, 188 (1999).
- E. N. Edinger, M. J. Risk, *Palaio* **9**, 576 (1994).
- E. Vennin et al., *Palaio* **21**, 163 (2004).
- F. R. Bosellini, C. Perrin, *Palaio* **25**, 71 (2008).
- J. W. Wells, in *Treatise on Marine Ecology and Paleocology*, J. W. Hedgpeth, Ed. (Geological Society of America, New York, 1957), vol. 67.
- P. Hallock, in *The History and Sedimentology of Ancient Reef Systems*, G. D. Stanley Jr., Ed. (Kluwer, New York, 2001), pp. 387–427.
- W. D. Allman, *Palaio* **16**, 9 (2001).
- E. N. Edinger, M. J. Risk, *Paleobiology* **21**, 200 (1995).
- A. von der Heydt, H. A. Dijkstra, *Paleoceanography* **21**, PA1011 (2006).
- A. O'Dea et al., *Proc. Natl. Acad. Sci. U.S.A.* **104**, 5501 (2007).
- S. Jain, L. S. Collins, *Mar. Micropaleontol.* **63**, 57 (2007).
- B. Schneider, A. Schmittner, *Earth Planet. Sci. Lett.* **246**, 367 (2006).
- J. B. C. Jackson, *Bioscience* **41**, 475 (1991).
- J. M. Pandolfi, J. B. C. Jackson, *Ecol. Lett.* **9**, 818 (2006).
- A. F. Budd, R. A. Petersen, D. F. McNeill, *Palaio* **13**, 170 (1998).
- L. E. Lisiecki, M. Raymo, *Paleoceanography* **20**, PA1003 (2005).
- J. B. C. Jackson, *Science* **264**, 1412 (1994).
- J. B. C. Jackson et al., *Science* **293**, 629 (2001).
- F. Schuster, *Proc. Ninth Int. Coral Reef Symp.* **1**, 199 (2002).
- We thank the innumerable field assistants, colleagues, and museum collections staff who helped us over the past 20 years to collect, curate, and identify the fossils that made this study possible. This research was supported by NSF grants EAR 9909485, EAR 0445789, and DBI 0237337; the Smithsonian Tropical Research Institute; and the William E. and Mary B. Ritter Chair at the Scripps Institution of Oceanography.

Supporting Online Material

www.sciencemag.org/cgi/content/full/319/5869/1521/DC1
Materials and Methods

Figs. S1 to S3

Tables S1 to S5

References

25 October 2007; accepted 11 February 2008

10.1126/science.1152197

Amyloid Fibrils of the HET-s(218–289) Prion Form a β Solenoid with a Triangular Hydrophobic Core

Christian Wasmer,* Adam Lange,* H el ene Van Melckebeke,* Ansgar B. Siemer,† Roland Riek, Beat H. Meier‡

Prion and nonprion forms of proteins are believed to differ solely in their three-dimensional structure, which is therefore of paramount importance for the prion function. However, no atomic-resolution structure of the fibrillar state that is likely infectious has been reported to date. We present a structural model based on solid-state nuclear magnetic resonance restraints for amyloid fibrils from the prion-forming domain (residues 218 to 289) of the HET-s protein from the filamentous fungus *Podospora anserina*. On the basis of 134 intra- and intermolecular experimental distance restraints, we find that HET-s(218–289) forms a left-handed β solenoid, with each molecule forming two helical windings, a compact hydrophobic core, at least 23 hydrogen bonds, three salt bridges, and two asparagine ladders. The structure is likely to have broad implications for understanding the infectious amyloid state.

Prions are infectious proteins capable of self-replicating their conformation and are best known as the agent of diseases such as scrapie in sheep (1), bovine spongiform encephalopathy in cattle (2), and a new variant of Creutzfeldt-Jakob disease in humans (3). Prions have also been described in yeast and filamentous fungi (4). The infectious form of prions has been characterized as a β sheet-rich molecular aggregate termed an amyloid fibril (5, 6).

HET-s is a protein of the filamentous fungus *Podospora anserina*. In its prion form, HET-s plays a role in heterokaryon incompatibility, a fungal self/nonself recognition phenomenon that prevents different forms of parasitism. The pro-

teinase K-resistant core of the prion fibrils formed by the C-terminal residues 218 to 289 is unstructured in solution and forms infectious fibrils in vitro (7). Earlier work showed that HET-s(218–289) in its fibrillar state consists of four β strands forming two windings of a β solenoid (8). However, it gave no information about the intermolecular β -sheet propagation (parallel or anti-

Physical Chemistry, ETH Zurich, 8093 Zurich, Switzerland.

*These authors contributed equally to this work.

†Present address: Department of Chemistry, Columbia University, New York, NY 10027, USA.

‡To whom correspondence should be addressed. E-mail: beme@ethz.ch

parallel), and the model was not based on distance restraints.

We used solid-state nuclear magnetic resonance (NMR) to determine the structure of the rigid core of HET-s(218–289) fibrils. Previous work on amyloid fibrils by NMR has shown that structural information can be obtained from these noncrystalline entities (9–14). Resonance assignment of HET-s(218–289) solid-state NMR spectra have been described previously (15, 16). We determine a large number of distance restraints from uniformly labeled samples and specifically identify purely intra- and intermolecular restraints by using differently labeled samples. The ^1H - ^1H and ^{13}C - ^{13}C internuclear distance restraints are derived from ^{13}C -detected proton-spin diffusion [carbon-proton-proton-carbon (CHHC) and nitrogen-proton-proton-carbon (NHHC) experiments (17, 18)] and from proton-driven ^{13}C spin diffusion (PDS) (19), respectively (for a list of experiments, see table S1). Part of a CHHC spectrum from a uniformly isotopically labeled (U - ^{13}C , ^{15}N) sample is shown in Fig. 1A as a representative example. From this spectrum, 41 structurally meaningful restraints were obtained (table S2b). These are translated to upper distance restraints of 3.5 and 4.5 Å, for strong and weak cross peaks, respectively. Sixty-two distance restraints (table S2b) were identified from a PDS spectrum of uniformly 2 - ^{13}C -labeled fibrils and translated to upper distance limits of 7.5 Å [2 - ^{13}C glycerol was used as a carbon source (20, 21) in this sample to reduce spectral overlap, line width, and relayed spin-diffusion effects].

In fibrils with uniform isotope labeling, distance restraints measured by NMR are difficult to assign to either intra- or intermolecular contacts. A comparison of spectra from uniformly labeled samples with those from “diluted samples,” in which isotopically labeled monomers are mixed with unlabeled material before fibrilization (ratio 1:2.5), can in some cases resolve the ambiguities. Following this strategy, we identified a total of 30 intramolecular restraints from a CHHC spectrum on a diluted U - ^{13}C , ^{15}N -labeled HET-s(218–289) sample and PDS spectra on diluted extensively labeled HET-s(218–289) (Fig. 1, B and C) (20, 21). Intermolecular restraints were obtained from a sample fibrilized from a mixture of U - ^{13}C -labeled and U - ^{15}N -labeled HET-s(218–289) molecules (22). Polarization transfer between ^{15}N - and ^{13}C -bound protons, respectively (17) (Fig. 1E), indicates short $\text{H}^{\text{N}}-\text{H}^{\text{C}}$ contacts (~ 3.0 Å) between β sheets of different monomers and thus selectively characterizes the intermolecular interface. Comparing these spectra to similar spectra on uniformly labeled compounds allowed us to identify seven H bonds as intramolecular and seven H bonds as intermolecular, and nine H bonds remained ambiguous (intra- or intermolecular). Details are given in the supporting online material (23). These H bonds define a parallel in-register arrangement of the intra- and intermolecular β -sheet interfaces, which is supported by 69 of the

experimental restraints that correlate residues i and $i + (36 \pm 1)$ (Fig. 1 and fig. S1). Fourteen of these can be identified unequivocally as intramolecular.

All restraints used for the structure calculation are summarized in Table 1 (a comprehensive list of cross peaks is given in table S2). In total, 90 ^{13}C - ^{13}C and 44 ^1H - ^1H distance restraints (i.e., 2.8 per assigned residue, all of them containing nontrivial structural information) were identified, together with the 23 β -sheet H bonds (fig. S2). In addition, 74 angle restraints obtained by TALOS (24) were used (fig. S2). From NMR (8) and mass-per-length measurements (25), it is known that the thinnest HET-s(218–289) fibril consists of a stack of single molecules all having the same structure. To implement the resulting quasi-one-dimensional (1D) symmetry, we used 206 additional intermolecular distance restraints (fig. S2). The NMR-structure calculation was performed with CYANA (26) using the restraints of Table 1 on a set of seven molecules, and yielded the structure depicted in Fig. 2. The backbone heavy-atom average root mean square deviation to the mean structure of the 20 lowest-energy conformers (200 calculated structures) is 0.4 Å for the backbone and 1.0 Å for all heavy atoms, considering only the rigid core of one HET-s(218–289) molecule (residues N226 to G242, N262 to G278) (Fig. 2 and fig. S3).

The overall organization of a HET-s(218–289) fibril is a left-handed β solenoid with two windings per molecule (Fig. 2A). The core of the fibril is defined by three β strands per winding (six β strands per molecule) that form continuous in-register parallel β sheets. An additional β sheet outside the core is formed by $\beta 2\text{b}$ and $\beta 4\text{b}$ (Fig. 2B). This organization is consistent with the fold proposed earlier (8), but the details show that each previously proposed β strand is split into two shorter segments (e.g., $\beta 1$ into $\beta 1\text{a}$ and $\beta 1\text{b}$). The segments $\beta 1\text{a}$ and $\beta 1\text{b}$ ($\beta 3\text{a}$ and $\beta 3\text{b}$) are connected by a two-residue β arc, changing the inside-outside pattern of side chains and leading to an approximately rectangular kink in the strand at K229 and E265, for $\beta 1$ and $\beta 3$, respectively (Fig. 2B). The connection between $\beta 1\text{b}$ and $\beta 2\text{a}$ (and similarly, between $\beta 3\text{b}$ and $\beta 4\text{a}$) is provided by a three-residue β arc, allowing for the orientation change of the polypeptide back-

bone by $\sim 150^\circ$. A disruption of the β -sheet pattern is also observed between $\beta 2\text{a}$ and $\beta 2\text{b}$ ($\beta 4\text{a}$ and $\beta 4\text{b}$), leading to $\sim 90^\circ$ β arcs. $\beta 1$ - $\beta 2$ and $\beta 3$ - $\beta 4$ are pseudo-repeats and form both parallel intramolecular and intermolecular H bonds, as follows: $\beta 1\text{a}$ - $\beta 3\text{a}$, $\beta 1\text{b}$ - $\beta 3\text{b}$, $\beta 2\text{a}$ - $\beta 4\text{a}$, $\beta 2\text{b}$ - $\beta 4\text{b}$ (Fig. 2, A and D).

As seen in Fig. 2, D to F, the β -sheet arrangement is stabilized by favorable side-chain contacts. The first three β strands of each pseudo-repeat enclose a triangular hydrophobic core that is tightly packed and contains almost exclusively hydrophobic residues (Ala, Leu, Ile, and Val) with numerous experimental restraints (indicated by green labels in the spectra of Fig. 1) between hydrophobic side chains. The packing is dense and defines a dry interface. The only polar residues in the core are T233 and S273, which can form side-chain H bonds, thereby stabilizing the formation of the turn between $\beta 1$ and $\beta 2$. In contrast, all charged residues face outside and are mostly located in β -arc regions, where the solvent accessibility is high. Three of them are arranged on top of each other such that charges are compensated (Fig. 2, D to F) and the formation of salt-bridges becomes possible. Several experimental restraints support the existence of the salt-bridges K229-E265, E234-K270, and R236-E272 (fig. S4). Because the stacking is parallel, we expect the charge compensation to have both intra- and intermolecular character. This may explain why fibrils have high stability against denaturation by non-ionic urea at neutral pH, but are destabilized by urea at acidic or basic pH (27). The two asparagines next to the hydrophobic core are stacked and can form a ladder (N226-N262), further contributing to the fibril stability through side-chain H bonds. Another asparagine ladder can be formed outside the hydrophobic core (N243-N279) (28).

For the structure calculation, only a limited number of well-resolved NMR peaks was used. To assess whether the structure is consistent with all the peaks observed in the spectra, we calculated the expected cross peaks from the internuclear distances in the structure. As an example, all expected resonances in the NHHC spectrum are shown as symbols in Fig. 1D. Similar good agreement was found for the other spectra (figs.

Table 1. Number of structural restraints used for the structure calculation (per monomer).

Type of restraint		Total	^1H - ^1H	^{13}C - ^{13}C	Unambiguously intramolecular
Distance restraints	Total	134	44	90	30
	Short-range ($ i - j \leq 1$)	12	12	—	—
	Medium-range ($2 \leq i - j \leq 4$)	9	2	7	3
	Register ($35 \leq i - j \leq 37$)	68	13	55	14
	Other long-range	45	17	28	13
		Total	Intramolecular	Intermolecular	Ambiguous
Hydrogen bond restraints		23	7	7	9
Periodicity restraints		206	—	206	—
Dihedral angle restraints (TALOS)		74	—	—	—

S5 and S6). In our structure, the cross-section of the fibril is approximately circular (Fig. 2C), a feature which might explain why no helical twist of HET-s(218–289) fibrils has been detected in electron micrographs (25). The structure given in Fig. 2 explains all the details of the chemical-shift data, H/D exchange, water accessibility, and mutant studies by Ritter *et al.* (8), even the ones that could not be explained by the earlier straight-stranded fold (8). The non- β -sheet chemical shift of K229 and E265 residues, together with their fast H/D exchange rates, can now be explained by the β arc at this position. Residues D230 and T266, which were found to be highly solvent-accessible as probed by chemical cross-linking of cysteine mutants (8), are indeed solvent-exposed

in our structure (Fig. 2, D to F), whereas I231, T233, V239, L241, V267, and V275 are buried in the hydrophobic core of the fibrils, explaining their very low water accessibility. G242, N279 and E280, for which intermediate values were found, are located in a region of β 2b and β 4b where the NMR structure is less well defined and where both sides of the β strand are somewhat protected (Fig. 2C). The present model also explains the requirement for a minimal length of the loop connecting β 2b and β 3a to maintain infectivity (8).

The structure of HET-s(218–289) shows an overall β -helical fold that is of higher structural complexity than that of peptide fibrils (29, 30)—a complexity that is reminiscent of soluble protein

folds. Part of the complexity, e.g., the favorable alternation of positive and negative charges in ladders along the fibril axis, can be realized only because of the pseudo-repeat (β 1/ β 3 and β 2/ β 4), leading to a structure in which one molecule forms two turns of the solenoid. (31). This feature distinguishes HET-s from other amyloids and prions that have also been modeled by β solenoids (32, 33). It leads to the formation of three salt-bridges that stabilize the structure, in contrast to the finding in yeast prion protein Sup35 where the presence of pseudo-repeats is probably related to structural variability and the existence of prion strains (31). The three-stranded triangular hydrophobic core indeed bears some resemblance to β -solenoid structures of soluble

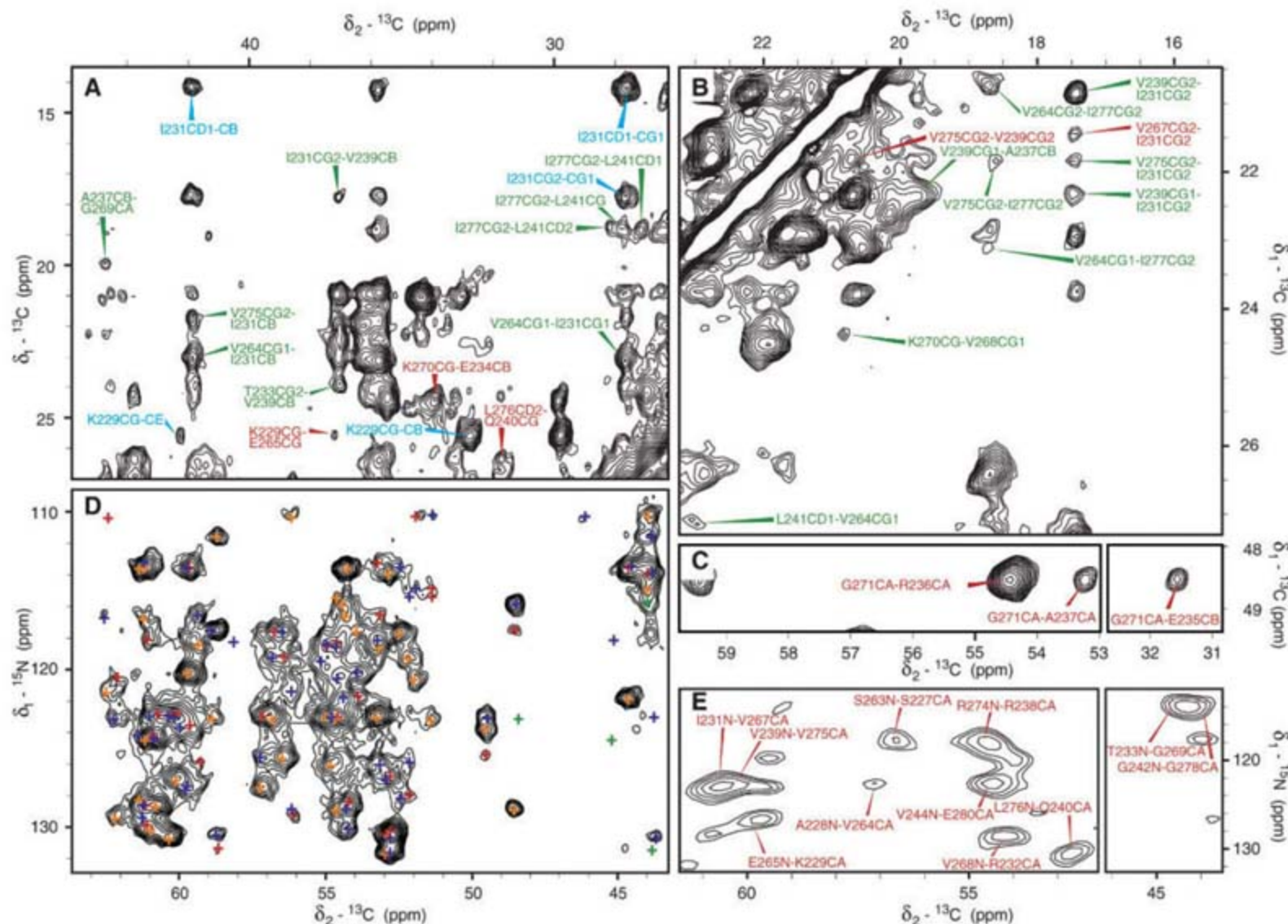
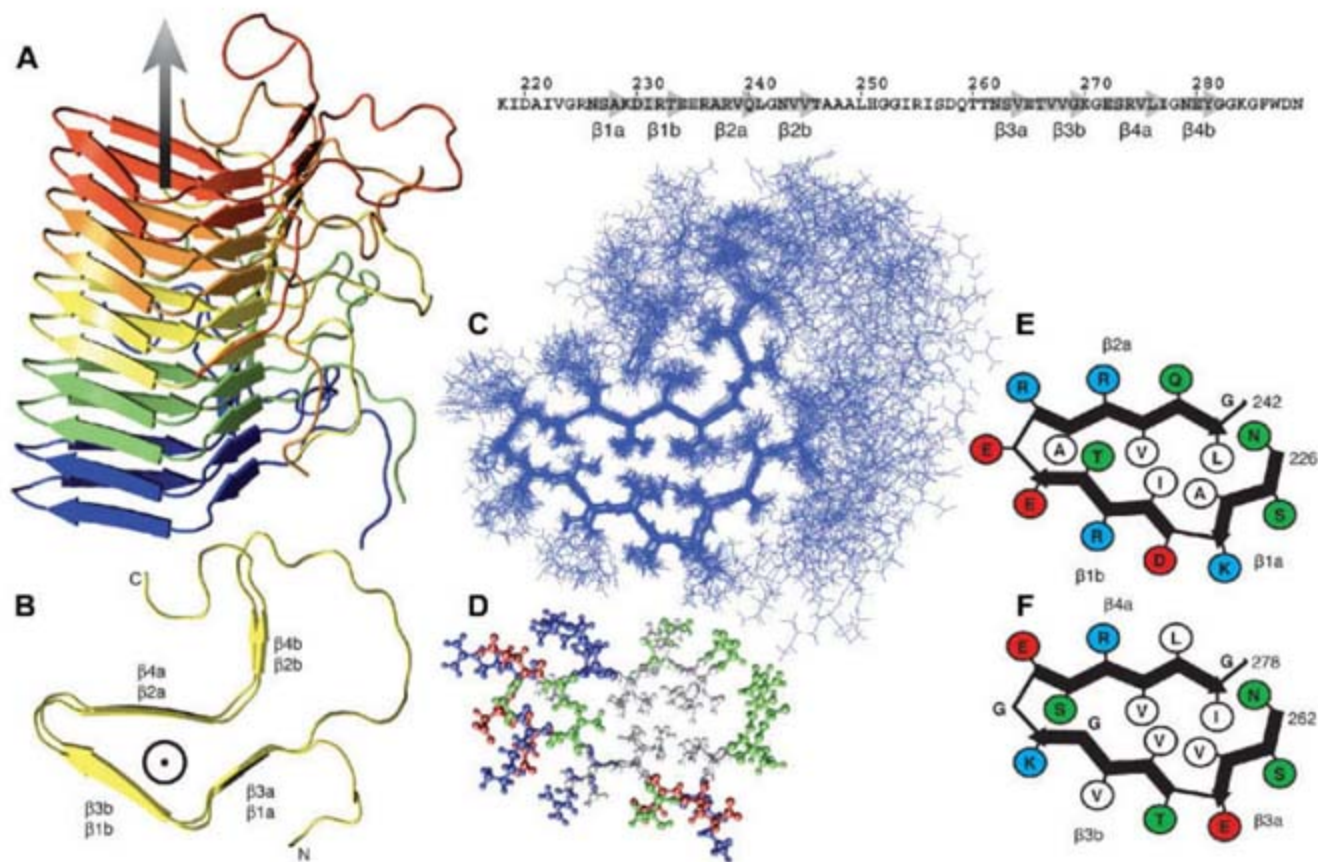


Fig. 1. Solid-state NMR spectra of HET-s(218–289) fibrils. Blue labels identify signals that correspond to short-range restraints ($|i - j| \leq 1$); red labels, signals that correspond to register-restraints ($35 \leq |i - j| \leq 37$); and green labels, restraints that define the hydrophobic core. (A) CHHC spectrum of U- ^{13}C , ^{15}N -labeled HET-s(218–289) fibrils recorded at 19-kHz magic-angle spinning (MAS) with 150- μs ^1H - ^1H mixing with register and side-chain contacts useful for structure calculation. (B) PDS spectrum of fibrils containing 25% 2 - ^{13}C -labeled monomers diluted in unlabeled HET-s(218–289), recorded at 10-kHz MAS with 250-ms ^{13}C - ^{13}C mixing giving intramolecular contacts. (C) PDS spectrum of fibrils containing 25% $1,3$ - ^{13}C -labeled molecules diluted in unlabeled HET-s(218–289), recorded at

19-kHz MAS with 500-ms ^{13}C - ^{13}C mixing giving further intramolecular contacts. (D) NHHC spectrum of a U- ^{13}C , ^{15}N -labeled sample recorded at 19-kHz MAS with 150- μs ^1H - ^1H mixing. The expected peak positions calculated for proton pairs separated by less than 3.5 Å in the lowest-energy structure are indicated by colored crosses. For a β sheet, three types of H_i^{N} - H_j^{α} contacts are mainly expected: intrasidue ($j = i$), sequential ($j = i - 1$), and interstrand register ($|j - i| = 36$, intra- and intermolecular) peaks are indicated with blue, orange, and red symbols, respectively. The green crosses indicate the positions of other expected peaks. (E) NHHC spectrum of fibrils containing a 1:1 mixture of U- ^{13}C - and U- ^{15}N -labeled monomers, recorded at 9.5-kHz MAS with 150- μs ^1H - ^1H mixing. Only intermolecular proton contacts give rise to peaks in this spectrum.

Fig. 2. Structure of the HET-s(218–289) fibrils. The fibril axis is indicated by an arrow. (A) Side view of the five central molecules of the lowest-energy structure of the HET-s(218–289) heptamer calculated from the NMR restraints. (B) Top view of the central molecule from (A). $\beta 3$ and $\beta 4$ lie on top of $\beta 1$ and $\beta 2$, respectively. A view orthogonal to the fibril axis is given in fig. S7. (C) NMR bundle: superposition on residues N226 to G242, N262 to G278 of the 20 lowest-energy structures of a total of 200 calculated HET-s(218–289) structures. Only the central molecule of the heptamer is shown. (D) Representation of the well-defined central core of the fibril (N226 to G242, N262 to G278). Hydrophobic residues are colored white, acidic residues red, basic residues blue, and others green (lowest-energy structure). (E and F) Schematic representations of the two windings in (D): the first winding [N226 to G242, displayed in (E)] of the β solenoid is located beneath the second one [N262



to G278, displayed in [F]). Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

proteins like filamentous hemagglutinin (34) and the P22 tailspike protein (35). In contrast to HET-s(218–289), these structures are not periodic, but the geometry of the triangular core is quite similar. Furthermore, a β -solenoid fold has also been proposed for the prion state of the human prion protein PrP on the basis of modeling and electron microscopy (36) and for the yeast prion Sup35 (37, 38).

The well-organized structure of the HET-s prion fibrils can explain the extraordinarily high order in these fibrils, as seen by NMR, as well as the absence of polymorphism caused by different underlying molecular structures at physiological pH conditions, because the specific nature of the interactions in the fibril excludes polymorphic molecular conformations with comparable stability. The fibril structure of HET-s(218–289) exemplifies the well-defined structure of a functional amyloid and illustrates the interactions that can stabilize their fold (39).

References and Notes

1. S. B. Prusiner, *Science* **216**, 136 (1982).
2. G. A. H. Wells et al., *Vet. Rec.* **121**, 419 (1987).
3. R. G. Will et al., *Lancet* **347**, 921 (1996).
4. L. Benkemoun, S. J. Saupé, *Fungal Genet. Biol.* **43**, 789 (2006).
5. U. Baxa, T. Cassese, A. V. Kajava, A. C. Steven, in *Advances in Protein Chemistry* (Academic Press, New York, 2006), vol. 73, pp. 125–180.
6. A. V. Kajava, A. C. Steven, in *Advances in Protein Chemistry* (Academic Press, New York, 2006), vol. 73, pp. 55–96.

7. A. Balguerie et al., *EMBO J.* **22**, 2071 (2003).
8. C. Ritter et al., *Nature* **435**, 844 (2005).
9. R. Tycko, *Protein Pept. Lett.* **13**, 229 (2006).
10. N. Ferguson et al., *Proc. Natl. Acad. Sci. U.S.A.* **103**, 16248 (2006).
11. C. P. Jaronec et al., *Proc. Natl. Acad. Sci. U.S.A.* **101**, 711 (2004).
12. H. Heise et al., *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15871 (2005).
13. F. Shewmaker, R. B. Wickner, R. Tycko, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 19754 (2006).
14. A. T. Petkova, W. M. Yau, R. Tycko, *Biochemistry* **45**, 498 (2006).
15. A. B. Siemer et al., *J. Biomol. NMR* **34**, 75 (2006).
16. A. B. Siemer et al., *J. Am. Chem. Soc.* **128**, 13224 (2006).
17. A. Lange, S. Luca, M. Baldus, *J. Am. Chem. Soc.* **124**, 9704 (2002).
18. A. Lange et al., *Angew. Chem. Int. Ed.* **44**, 2089 (2005).
19. N. M. Szeverenyi, M. J. Sullivan, G. E. Maciel, *J. Magn. Reson.* **47**, 462 (1982).
20. F. Castellani et al., *Nature* **420**, 98 (2002).
21. D. M. LeMaster, D. M. Kushlan, *J. Am. Chem. Soc.* **118**, 9255 (1996).
22. M. Etzkorn, A. Böckmann, A. Lange, M. Baldus, *J. Am. Chem. Soc.* **126**, 14746 (2004).
23. Materials and methods are available as supporting material on Science Online.
24. G. Cornilescu, F. Delaglio, A. Bax, *J. Biomol. NMR* **13**, 289 (1999).
25. A. Sen et al., *J. Biol. Chem.* **282**, 5545 (2007).
26. P. Guntert, C. Mumenthaler, K. Wuthrich, *J. Mol. Biol.* **273**, 583 (1997).
27. R. Sabate et al., *J. Mol. Biol.* **370**, 768 (2007).
28. M. D. Yoder, N. T. Keen, F. Jurnak, *Science* **260**, 1503 (1993).
29. R. Nelson et al., *Nature* **435**, 773 (2005).
30. M. R. Sawaya et al., *Nature* **447**, 453 (2007).

31. B. H. Toyama, M. J. S. Kelly, J. D. Gross, J. S. Weissman, *Nature* **449**, 233 (2007).
32. A. V. Kajava, J. M. Squire, D. A. D. Parry, in *Advances in Protein Chemistry* (Academic Press, New York, 2006), vol. 73, pp. 1–15.
33. N. D. Lazo, D. T. Downing, *Biochemistry* **37**, 1731 (1998).
34. B. Clantin et al., *Proc. Natl. Acad. Sci. U.S.A.* **101**, 6194 (2004).
35. S. Steinbacher et al., *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10584 (1996).
36. C. Govaerts, H. Wille, S. B. Prusiner, F. E. Cohen, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 8342 (2004).
37. R. Krishnan, S. L. Lindquist, *Nature* **435**, 765 (2005).
38. A. Kishimoto et al., *Biochem. Biophys. Res. Commun.* **315**, 739 (2004).
39. D. M. Fowler, A. V. Koulouf, W. E. Balch, J. W. Kelly, *Trends Biochem. Sci.* **32**, 217 (2007).
40. We thank C. Ritter for support in producing the samples, S. Saupé for scientific discussions, and the Swiss National Science Foundation and the TH-system of the ETH Zurich for financial support. H.V.M. acknowledges a stipend by the European Union (Marie Curie EIF, FP6) and A.L. by the European Molecular Biology Organization (EMBO). The bundle of 20 conformers representing the NMR structure has been deposited in the Protein Data Bank with accession code 2RNM, and assignments have been deposited in the Biological Magnetic Resonance Data Bank (ID 11028).

Supporting Online Material

www.sciencemag.org/cgi/content/full/319/5869/1523/DC1
 Materials and Methods
 Figs. S1 to S7
 Tables S1 to S3
 References

17 October 2007; accepted 6 February 2008
 10.1126/science.1151839

A Retrotransposon-Mediated Gene Duplication Underlies Morphological Variation of Tomato Fruit

Han Xiao,¹ Ning Jiang,² Erin Schaffner,^{1,3*} Eric J. Stockinger,¹ Esther van der Knaap^{1†}

Edible fruits, such as that of the tomato plant and other vegetable crops, are markedly diverse in shape and size. *SUN*, one of the major genes controlling the elongated fruit shape of tomato, was positionally cloned and found to encode a member of the IQ67 domain-containing family. We show that the locus arose as a result of an unusual 24.7-kilobase gene duplication event mediated by the long terminal repeat retrotransposon *Rider*. This event resulted in a new genomic context that increased *SUN* expression relative to that of the ancestral copy, culminating in an elongated fruit shape. Our discovery demonstrates that retrotransposons may be a major driving force in genome evolution and gene duplication, resulting in phenotypic change in plants.

Most crop species display distinctly different morphologies in comparison with those of the relatives from which they were domesticated (1, 2). For tomato (*Solanum lycopersicum*), cultivated types bear fruit of varying sizes with many diverse shapes (2). The locus *sun* comprises a major quantitative trait locus explaining 58% of the phenotypic variation in a cross that was derived from the *S. lycopersicum* variety "Sun1642," which has elongated shaped fruit, and its wild relative *S. pimpinellifolium* (accession LA1589), which has round fruit (3). Fine mapping indicates that *sun* resides in a region of the tomato genome on the short arm of chromosome 7 that carries small-scale insertions, deletions, and tandem duplications (4). One insertion, estimated to be 30 kb, is present in Sun1642 (but not in LA1589) and is linked to fruit shape (4).

We constructed a genetic and physical map of the *sun* locus in both parents [supporting online material (SOM) Materials and Methods and fig. S1, A and B]. Comparative sequence analysis of the locus showed that the size difference was due to the insertion of a 24.7-kb segment present in Sun1642 but absent from LA1589 (Fig. 1A). This insertion completely coassorted with the phenotype (fig. S1B), implying that it underlies the molecular basis of elongated fruit shape. A Southern blot showed that LA1589 contained only one copy of this sequence, whereas Sun1642 contained two copies (fig. S1C). Both accessions shared a copy of this locus on chromosome 10, suggesting that the inserted segment on chromosome 7 in Sun1642 originated from chromosome 10 (Fig. 1B and fig. S1C). Comparative sequence analysis of the 24.7-kb duplicates showed that *IQD12* was rearranged relative to the other se-

quences within the 24.7-kb segment (Fig. 1). Moreover, the only nucleotide divergence between the two copies was a 3-base pair (bp) mismatch present at the breakpoint of the rearrangement, which suggests a recent origin of the duplication. Close inspection of the sequences indicated that a *Copia*-like retrotransposon, which we named *Rider*, was present at both loci. *Rider* resembled a long terminal repeat (LTR) retrotransposon because it was flanked by identical 398-bp LTRs (designated in red as "1" and "2") and a 5-bp target site duplication (TSD) sequence GACCT (Fig. 1B). Similarly, the chromosome 7 region had the identical *Rider* element flanked by two LTRs (LTR1 and LTR2). One additional LTR (LTR3) was identified further downstream of *Rider* on chromosome 7, which (together with LTR1) flanked the entire duplicated fragment (Fig. 1A). At the site of the presumed integration, immediately flanking LTR1 and LTR3, a 5-bp motif ATATT was identified, resembling a TSD of a trans-

position event (Fig. 1A). Sequencing of LA1589 supported that the integration of the segment occurred at this 5-bp motif. These data, along with analyses of polypurine tracts (PPT) (SOM Text), suggest that the entire 24.7-kb fragment transposed via *Rider* from chromosome 10 to chromosome 7 as one large retrotransposon.

Because the fruit-shape phenotype at *sun* was entirely linked to the inserted segment (fig. S1B), we considered that the elongated fruit shape may be due to genes present on the transposed fragment or the disruption of a gene as a result of the insertion event. We identified five putative genes at the *sun* locus in Sun1642 as well as *Rider* (Fig. 1A). Four of these genes were *IQD12* [which belonged to the IQ67 domain-containing plant proteins (5)], an *SDL1*-like gene with high similarity to a tobacco and *Arabidopsis* gene (6, 7), and two hypothetical genes *HYP1* and *HYP2* encoding a protein with weak similarity to CUC1 (8) and a 487-amino acid protein with high similarity to a potato protein of unknown function (AY737314), respectively. The fifth gene was *DEFL1*, encoding a putative secreted defensin protein (BT012682). The insertion into the intron of this gene strongly suggested that *DEFL1* in Sun1642 was disrupted (Fig. 1A). The *Rider* retrotransposon contained a single open reading frame, which encoded a 1307-amino acid polyprotein including integrase and reverse transcriptase.

To test the effect of the *sun* locus on fruit shape in a homogeneous background and to ascertain whether a difference in expression of one of the candidate genes might underlie the phenotype, we generated a set of near-isogenic lines (NILs) that differed at *sun*. The NIL in the LA1589 background showed negligible ovary-shape differences in floral buds 10 days before flower opening (Fig. 2A) ($n = 10$ buds, $P = 0.64$). At anthesis, ovary shapes began to show sig-

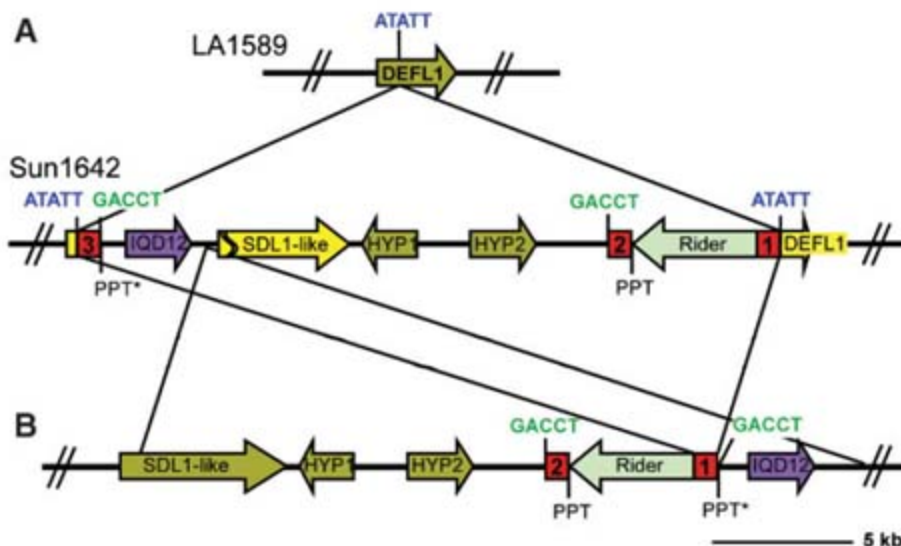


Fig. 1. Structural organization of the *sun* locus in LA1589 and Sun1642 (A) and the ancestral locus on chromosome 10 (B). Arrows show directionality of the predicted genes and pseudogenes. Dark green arrows indicate ab initio predicted genes, purple arrows indicate the rearranged *IQD12* gene, light green arrows indicate the LTR retrotransposon *Rider*, and yellow arrows indicate pseudogenes. Red numbered boxes identify *Rider*'s LTRs and are numbered according to the predicted order of transcription. The TSD sequences are indicated in green (chromosome 10) and blue (chromosome 7). The position of the original PPT and predicted actual PPT* are indicated (see SOM Text for more details).

¹Department of Horticulture and Crop Science, Ohio State University/Ohio Agricultural Research and Development Center, Wooster, OH 44691, USA. ²Department of Horticulture, Michigan State University, East Lansing, MI 48824, USA. ³College of Wooster, Wooster, OH 44691, USA.

*Present address: College of Medicine, University of Toledo, Toledo, OH 43606, USA.

†To whom correspondence should be addressed. E-mail: vanderknaap.1@osu.edu

nificant, albeit slight, differences ($n = 40$ ovaries, $P = 0.04$). The most significant differences in shape were found in developing fruit 5 days after

anthesis (Fig. 2A) ($n = 35$ fruits, $P < 0.0001$), indicating that shape change primarily followed pollination and fertilization (3).

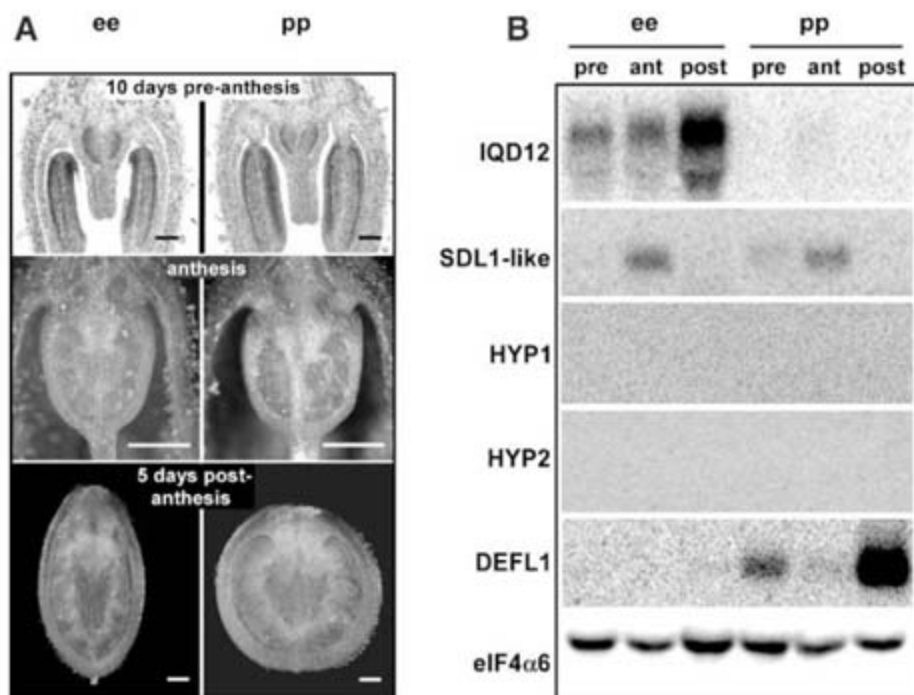
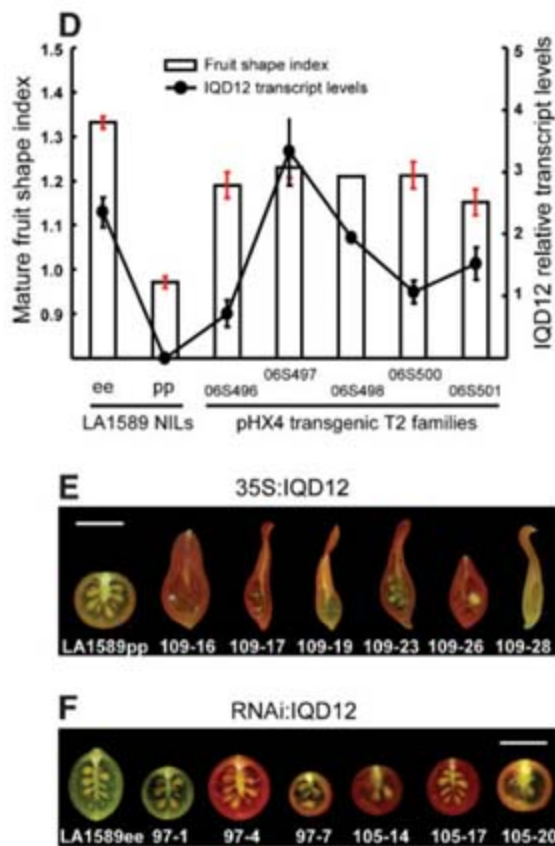
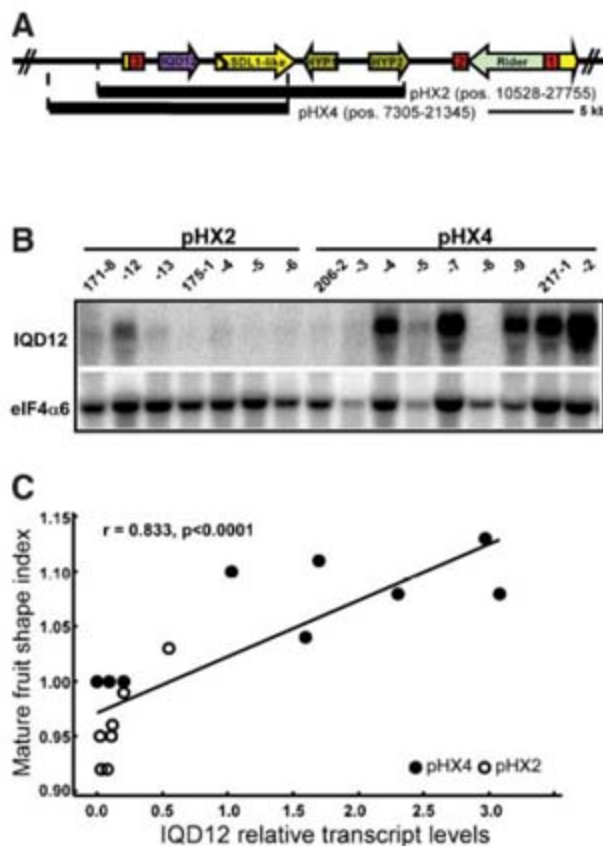


Fig. 2. *sun* affects fruit shape after anthesis. Data shown are from NILs in the LA1589 background. “ee” denotes homozygous Sun1642, whereas “pp” denotes homozygous LA1589 “pre,” pre-anthesis; “ant,” -anthesis; “post,” post-anthesis. (A) Representative ovaries at three developmental stages. Scale bars for each stage (from top to bottom) are 100, 500, and 500 μ m, respectively. (B) Expression analysis of five candidate genes at *sun*. Total RNA was isolated from tissues of each genotype at the same developmental stages shown in (A), blotted, and hybridized with the probes indicated.

A much higher transcript level of *IQD12* was observed in the NIL carrying the Sun1642 allele in comparison with the NIL carrying the LA1589 allele (Fig. 2B). The highest transcript levels of *IQD12* were found in young developing fruit 5 days after anthesis. Transcript levels of the disrupted gene *DEFL1* showed an inverse expression pattern relative to *IQD12*: When *IQD12* was expressed, *DEFL1* was not (and vice versa) (Fig. 2B). Reverse transcription polymerase chain reaction failed to detect any *DEFL1* transcript in NILs that carried the chromosomal duplication, indicating that *DEFL1* expression was abolished as a result of *Rider*’s transposition into this gene (Fig. 1). Transcript levels of *SDL1-like*, *HYP1*, and *HYP2* were not altered in the NILs, or were undetectable, and thus were deemed less likely candidates for the *SUN* gene.

Fruit-shape phenotype was hypothesized to be dosage-dependent, because heterozygous *sun* NIL plants exhibit an intermediate fruit shape relative to that of the parents in both the LA1589 and Sun1642 backgrounds (fig. S2). Northern blots demonstrated that *IQD12* was expressed about twofold higher in individuals homozygous for the transposed fragment than in heterozygous plants. Similarly, *DEFL1* was expressed about twofold higher in homozygous individuals lacking the transposed fragment than in heterozygotes (fig. S2). Thus, the transposition event most likely placed *IQD12* under the cis-regulatory control of factors normally conferring

Fig. 3. The regulation of tomato fruit shape is controlled by increased transcription of *IQD12*. Data shown are in the LA1589 background. (A) Graphical depiction of the two genomic constructs used to complement *sun*. Features of the genomic region are described in Fig. 1. (B) Northern blot analysis of T1 plants transformed with pHX2 and pHX4 constructs with *IQD12* as probe with RNA isolated from fruit 5 days after anthesis. (C) Fruit-shape index of T1 lines significantly correlates with *IQD12* transcript levels. (D) Fruit-shape index and *IQD12* transcript levels of homozygous T2 plants derived from five independent pHX4 primary transformants. The average fruit-shape index (column) and *IQD12* transcript levels (line) were determined from the Northern blot data shown in (fig. S2B). Error bars denote SD. For comparison, fruit-shape index and *IQD12* transcript levels of the NIL were included. (E) Constitutive expression of *IQD12* in the round-fruited LA1589 background results in extremely elongated fruit. Each fruit represents an independently transformed plant. Fruit from the non-transformed round-fruited NIL (LA1589pp) is shown for comparison. (F) RNAi-mediated reduced expression of *IQD12* in the NIL carrying elongated



fruit (LA1589ee) results in a notable reduction in fruit elongation. Each fruit represents an independently transformed plant. The fruit from the nontransformed NIL (LA1589ee) is shown for comparison. Scale bars in (E) and (F) represent 1 cm.

high levels of *DEFL1* expression in developing fruit.

The insert from two overlapping Sun1642 λ genomic clones was transformed into LA1589 and the round-fruited NIL in the Sun1642 background. These clones—pHX2 and pHX4—harbored the full-length *IQD12* gene, including the promoter shared with the ancestral gene and the LTR, but contained different 5' as well as 3' end points (Fig. 3A). In the LA1589 background, most plants transformed with the pHX4 construct expressed *IQD12* at high levels, whereas this gene was expressed at low levels in the pHX2 transformants (Fig. 3B and table S1). Moreover, pHX4 T1 plants exhibited a greater fruit-shape index (ratio of height to width), whereas those with the pHX2 construct did not. Regression analysis confirmed that there was a significant linear relationship between *IQD12* transcript levels and fruit-shape index, which was in turn correlated to transformation with the pHX4 construct (Fig. 3C). Transformation of the same constructs into the round-fruited NIL in the Sun1642 background produced similar results (table S2).

Increased expression of *IQD12* and increased fruit-shape index cosegregated with the presence of the pHX4 construct in the next generation (Fig. 3D and fig. S3A). Transcript levels of *IQD12* and fruit-shape index of homozygous pHX4 plants indicated that, for most transgenic plant

families, the fruit-shape index and the *IQD12* transcript levels were nearly restored to the endogenous levels displayed by the NIL carrying the transposed segment (Fig. 3D, fig. S3, and table S1). Thus, control of fruit shape in tomato is regulated at the level of transcription of *IQD12*. In addition, important cis-element(s) in the 3.2-kb region upstream of *DEFL1* drive *IQD12* expression in the developing fruit, because this region was present in pHX4 and absent from pHX2 (Fig. 3, A to C). Additionally, the *SDL1-like* and *HYP1* genes probably do not affect fruit shape, because these genes were present on the pHX2 construct, which did not complement the phenotype (Fig. 3, A to C).

Although transcript levels and fruit-shape index were substantially increased in the transgenic lines carrying pHX4, they were not entirely restored to the levels of that in the NIL carrying the transposed segment (table S1). Thus, either additional cis-elements in the *DEFL1* 5' region or other genes contribute to fruit-shape phenotype (Fig. 3D and fig. S3). To determine whether *IQD12* alone was sufficient to confer an elongated shape to tomato, we overexpressed this gene in LA1589 and used RNA interference (RNAi) to reduce expression in plants carrying the transposed segment exhibiting elongated fruit. Six of 13 LA1589 lines transformed with *IQD12* under control of the constitutive cauliflower mosaic virus (CaMV) 35S RNA promoter bore extremely

elongated fruit and expressed *IQD12* at high levels (Fig. 3E, fig. S4A, and table S3), which cosegregated with the presence of the transgene in the next generation (fig. S4B and table S3). An *IQD12* inverted repeat construct—transformed into NILs carrying the transposed segment—resulted in rounder fruit similar to that found in the wild type, which was associated with decreased *IQD12* expression (Fig. 3E, fig. S4, C to E, and table S3). Thus, increased and reduced expression of *IQD12* confers changes to tomato fruit shape, which implies that this gene is necessary and sufficient in controlling shape at the *sun* locus. Hence, *IQD12* was renamed *SUN*.

All members of the IQD family contain the 67-amino acid IQ67 domain, which has only been found in plants (5). AtIQD1 binds calmodulin and is localized in the nucleus, and its overexpression increases glucosinolate levels in *Arabidopsis* (9). On the basis of the shared domain structure of the entire protein and phylogenetic analysis of the IQ67 motif, *SUN* is most closely related to AtIQD12, although the criterion of conservation of intron position suggests that it is more closely related to AtIQD11 (fig. S5A). Both *Arabidopsis* genes are members of the subfamily II of the IQD family of proteins that also include AtIQD13 and AtIQD14 (5) (fig. S5B). Extensive database searches identified members of this family from other dicot species and suggested that the IQD cluster containing *SUN* expanded to at least four members in tomato and potato, after the split from *Arabidopsis* (fig. S5B).

We propose that transcription of *Rider* on chromosome 10 began in LTR1. Instead of stopping in LTR2, readthrough occurred in the flanking genomic region (Fig. 4A, SOM Text, and fig. S7). In the first intron of the *SDL1-like* gene, the RNA polymerase switched to a region 3' of *SUN*. The 5-bp direct repeat GCAGA on each side of the breakpoint suggests that a template switch occurred (Fig. 4A), resulting in an elongated transcript that terminated in LTR1 (Fig. 4B). The introns of the *SUN* and *SDL1-like* genes were retained in the antisense orientation with regard to *Rider* and were not recognized and removed by the splicing machinery after RNA synthesis. This large element inserted into chromosome 7 (Fig. 4C), resulting in the gene arrangement at the *sun* locus (Fig. 1A). *Rider*'s 3' gene transduction and template switch are essential for a successful transposition, because two flanking LTRs are required for reverse transcription and double-stranded DNA synthesis of LTR retrotransposons (10). In addition, this model demonstrates that LTR retrotransposons can facilitate gene duplications, resulting in phenotypic change.

With the exception of AtIQD1 and *SUN*, the function of other IQ67 domain-containing proteins remains unknown. Unlike AtIQD1, *SUN* does not appear to be affecting glucosinolate levels in tomato, because these metabolites are not produced in solanaceous plants. However, AtIQD1 plays a role in the transcript accumulation of a small group of cytochrome P450 genes including

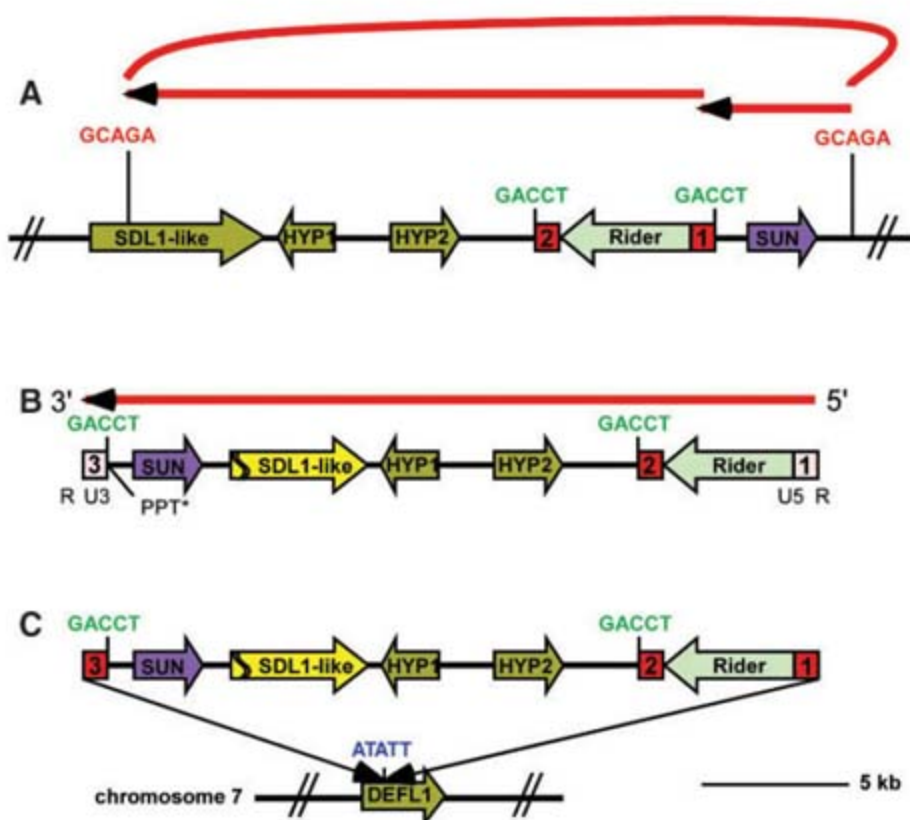


Fig. 4. A model for the segmental duplication and rearrangement at the *sun* locus. The features of the genomic region are described in Fig. 1. (A) Readthrough transcription of *Rider* on chromosome 10 to the signature sequence GCAGA in the first intron of *SDL1-like*. RNA polymerase switched the template to continue mRNA transcription from an identical signature sequence (GCAGA), located downstream of *SUN*. (B) Formation of one large retrotransposon mRNA. (C) After reverse transcription and second-strand cDNA synthesis, *Rider* integrated into chromosome 7 at the ATATT site located in the intron of *DEFL1*.

CYP79B3 and *CYP79B2* (9, 11). The encoded proteins catalyze the conversion of tryptophan into indole-3-acetaldoxime in tryptophan-dependent auxin biosynthesis in *Arabidopsis* (12–14). In tomato, overexpression of *SUN* resulted in extremely elongated and often seedless fruit, reminiscent of the parthenocarpic, elongated, and pointed tomato fruit resulting from expression of the auxin biosynthesis gene *iaaM* controlled by the placenta and ovule-specific *DeffH9* promoter (15, 16). The extremely elongated fruit shape and lack of proper seed development when *SUN* is overexpressed, in addition to its potential biochemical function, suggest that *SUN* may affect auxin levels or distribution in the fruit (Fig. 3E). It is therefore plausible that involvement of *SUN* in shape variation is through regulation of plant hormone and/or secondary metabolite levels, thereby affecting the patterning of the fruit.

References and Notes

- J. F. Doebley, B. S. Gaut, B. D. Smith, *Cell* **127**, 1309 (2006).
- I. Paran, E. van der Knaap, *J. Exp. Bot.* **58**, 3841 (2007).
- E. van der Knaap, S. D. Tanksley, *Theor. Appl. Genet.* **103**, 353 (2001).
- E. van der Knaap, A. Sanyal, S. A. Jackson, S. D. Tanksley, *Genetics* **168**, 2127 (2004).
- S. Abel, T. Savchenko, M. Levy, *BMC Evol. Biol.* **5**, 72 (2005).
- K. Lertpiriyapong, Z. R. Sung, *Plant Mol. Biol.* **53**, 581 (2003).
- A. Majira, M. Domin, O. Grandjean, K. Gofron, N. Houba-Herlin, *Plant Mol. Biol.* **50**, 551 (2002).
- S. Takada, K. Hibara, T. Ishida, M. Tasaka, *Development* **128**, 1127 (2001).
- M. Levy, Q. Wang, R. Kaspi, M. P. Parrella, S. Abel, *Plant J.* **43**, 79 (2005).
- B. Lewin, *Genes IX* (Jones and Bartlett, Sudbury, MA, ed. 9, 2008).
- C. D. Grubb, S. Abel, *Trends Plant Sci.* **11**, 89 (2006).
- A. K. Hull, R. Vij, J. L. Celenza, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2379 (2000).
- M. D. Mikkelsen, C. H. Hansen, U. Wittstock, B. A. Halkier, *J. Biol. Chem.* **275**, 33712 (2000).

- Y. Zhao *et al.*, *Genes Dev.* **16**, 3100 (2002).
- N. Ficcadenti *et al.*, *Mol. Breed.* **5**, 463 (1999).
- T. Pandolfini, G. L. Rotino, S. Camerini, R. Defez, A. Spena, *BMC Biotechnol.* **2**, 1 (2002).
- We thank T. Meulia and A. Kaszas at the Molecular and Cellular Imaging Center in Wooster for microscopy and sequencing, as well as D. Francis, S. Kamoun, D. Mackey, and P. Springer for comments. This work was funded in part by NSF: DBI 0227541 and DBI 0400811 (to E.v.d.K.); DBI 0110124 (to E.J.S.). Accession numbers of deposited sequences: EF094939, EF094940, EF094941, and EU491503.

Supporting Online Material

www.sciencemag.org/cgi/content/full/319/5869/1527/DC1

Materials and Methods

SOM Text

Figs. S1 to S7

Tables S1 to S4

References

15 November 2007; accepted 8 February 2008

10.1126/science.1153040

Evidence for Karyogamy and Exchange of Genetic Material in the Binucleate Intestinal Parasite *Giardia intestinalis*

Marianne K. Poxleitner,^{1*} Meredith L. Carpenter,^{1*} Joel J. Mancuso,¹ Chung-Ju R. Wang,¹ Scott C. Dawson,² W. Zacheus Cande^{1†}

The diplomonad parasite *Giardia intestinalis* contains two functionally equivalent nuclei that are inherited independently during mitosis. Although presumed to be asexual, *Giardia* has low levels of allelic heterozygosity, indicating that the two nuclear genomes may exchange genetic material. Fluorescence in situ hybridization performed with probes to an episomal plasmid suggests that plasmids are transferred between nuclei in the cyst, and transmission electron micrographs demonstrate fusion between cyst nuclei. Green fluorescent protein fusions of giardial homologs of meiosis-specific genes localized to the nuclei of cysts, but not the vegetative trophozoite. These data suggest that the fusion of nuclei, or karyogamy, and subsequently somatic homologous recombination facilitated by the meiosis gene homologs, occur in the giardial cyst.

Giardia intestinalis (syn. *lamblia*, *duodenalis*) is a common intestinal protozoan parasite and a major cause of water-borne diarrheal disease worldwide (1). As a diplomonad, it has two apparently genetically identical, functionally equivalent diploid nuclei (2–4), and the sequenced genome shows little heterozygosity (5, 6). Both nuclei are transcriptionally active and contain two complete copies of the genome (2–4). The parasite has two developmental stages: a binucleate, flagellated trophozoite that attaches to the upper small intestine, and the quadrinucleate infectious cyst, which is excreted from the body and persists in the water supply (7). In the encysting trophozoite, the two nuclei are thought to

undergo a nuclear division without a subsequent cytokinesis, giving rise to a mature cyst with four nuclei (8). Upon excystation, the cell goes through one cellular division and then one nuclear and cellular division to form four binucleate trophozoites (8).

The two nuclei in the trophozoite remain physically and genetically distinct during mitosis, with two autonomous spindles segregating the parental nuclei into the daughter cells (9). Neither mating nor meiosis has been reported in *Giardia*. However, if *Giardia* is asexual, it should accumulate substantial allelic heterozygosity within and between the two nuclei, and this has not been observed (5, 6, 10).

To determine whether the cyst nuclei remain physically autonomous, as they do in trophozoites, or whether nuclei can exchange genetic material, we performed fluorescent in situ hybridization (FISH) on trophozoites and cysts (9). In *Giardia*, stably transfected plasmids (episomes) are found in only one of the two nuclei

of trophozoites (Fig. 1A). This pattern persists throughout cell division and cytokinesis (9, 11). When we use FISH to detect episomes in cysts, we see two distinct patterns ($n = 50$): 72% of the cysts have several episomes present in two of the four nuclei (Fig. 1B), whereas 28% of the cysts have episomes in three of the four nuclei (Fig. 1C). This result suggests that in roughly one-third of encysting cells, two nuclei come together and their nuclear envelopes fuse, resulting in plasmid transfer between the nuclei.

To ascertain whether nuclear fusion occurs in *Giardia* cysts, we performed transmission electron microscopy (TEM). Cysts were prepared for TEM with high-pressure rapid freezing–freeze substitution to achieve optimal cytological preservation and eliminate chemical fixation artifacts due to the impermeability of the cyst wall. Using this method, we observed nuclei joined by contiguous membranes (Fig. 1, D and E). These nuclear envelope (NE) fusion morphologies are not derived from an incomplete mitosis because there are no spindle microtubules associated with the NE, and the multiple axonemes associated with paired and fusing nuclei differ in location from the axoneme pairs found at each spindle pole in mitotic cells (9). Of the ~100 individual cysts examined, we observed 15 with contiguous NEs. Because these cells were collected at a single moment in time, the percentage of nuclear fusions is less than that seen by FISH, which represents the net outcome of nuclear fusions over time. Thus, the nuclei in the *Giardia* cyst fuse (undergo karyogamy), and this fusion likely facilitates the transfer of episomes between nuclei.

We next sought to determine when karyogamy takes place in the encystation process. We counted the number of nuclei in cysts and the total number of cysts present at six different time points (Fig. 2A). Cysts appeared 12 hours after induction of encystation; subsequently, multiple nuclear arrangements were seen in cysts at every time point, suggesting that encystation is not well synchronized and that cells at varying stages of

¹Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA. ²Section of Microbiology, 255 Briggs Hall, One Shields Avenue, University of California, Davis, CA 95616, USA.

*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: zcande@berkeley.edu

encystation are always present. Three nuclear arrangements could be distinguished at intermediate time points: two pairs of nuclei at opposite ends of the cell, two pairs of nuclei close together, and three nuclei (a pair of nuclei and one large chromatin mass) usually close together (Fig. 2A).

Cysts with four nuclei close together were always the most abundant. These data are consistent with the hypothesis that at the four-nuclear stage, the nuclear pairs migrate from opposite ends of the cell to the same quadrant, undergo fusion, and then separate again.

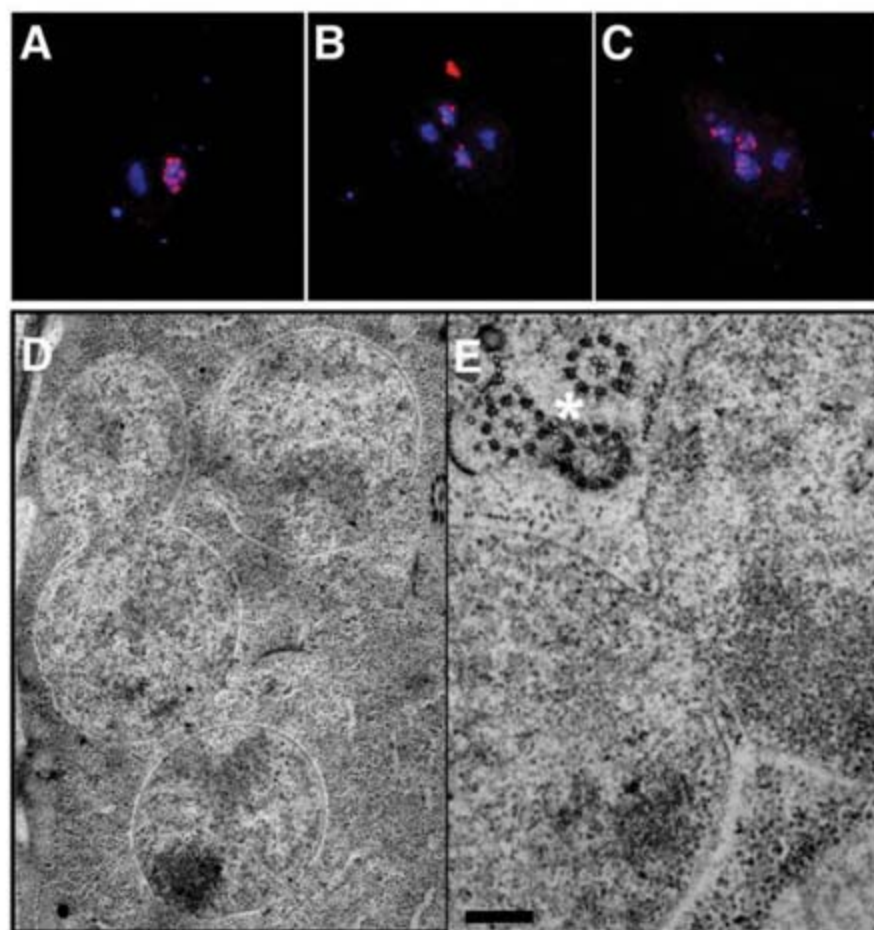
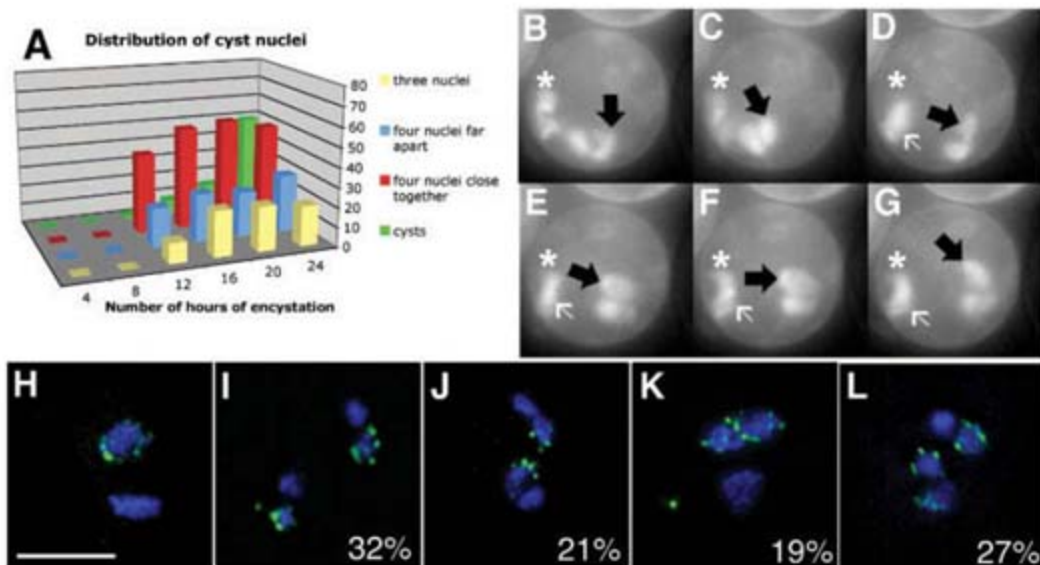


Fig. 1. Plasmids detected by FISH in *Giardia* trophozoites and cysts, and TEM showing nuclear fusion (karyogamy) in cysts. (A) When FISH is performed on trophozoites containing transformed plasmids, plasmid probes (red spots) are detected in only one of the two 4', 6'-diamidino-2-phenylindole (DAPI)-stained (blue) nuclei. (B and C) In cysts, plasmid probes are seen in either two of the four nuclei (B) or three of the four nuclei (C). (D) TEM image of a mature cyst. All four cyst nuclei are in the same plane; two of the nuclei are fused with contiguous nuclear envelopes. (E) Two fused nuclei from a different cell. Cytoplasmic axonemes are marked with an asterisk. Bar, 200 nm.

Fig. 2. Nuclear movements and plasmid distribution in *Giardia* cysts. (A) Time course of nuclear number and position during encystation. (B to G) Selected frames 8 s apart from a movie of a live, acridine orange-stained cyst. The black arrow follows the movement of the nuclear pair on the right side of the cyst. The nuclei within each pair also oscillate [asterisk in (B) to (G)]. The possible fusion of the left pair of nuclei is indicated by a white arrow in (D) to (G). Also see movies S1 to S3 (11). (H to L) Nuclear configuration and plasmid distribution. (H) In trophozoites, the plasmids (green spots) are detected by FISH in only one of the two nuclei. (I) Plasmids are detected in two of the four cyst nuclei when the two nuclear pairs are far apart. (J) Plasmids are detected in two of the four nuclei and the two nuclear pairs are close together. (K) Cyst nuclei have presumably undergone fusion, and plasmids are detected in two of the three nuclei. (L) Plasmids are present in three of the four nuclei. Percentages of each class of cyst are indicated in (I) to (L). Bar, 5 μ m.



Time-lapse images of live, acridine orange-stained cysts illustrate this dynamic nuclear behavior (Fig. 2, B to G and movies S1 to 3). Each nuclear pair exhibits two forms of movement (Fig. 2, B to G). The pair on the right moves relative to the pair on the left, and in 32 s rotates from a horizontal (Fig. 2B) to a vertical (Fig. 2F) orientation. The nuclei within each pair also fluctuate back and forth (Fig. 2, B to G). The nuclear movements in cysts are very different from what is observed in trophozoites, in which the two nuclei are tethered to the cytoskeleton and exhibit no movement (movie S2) except during mitosis. The dynamic movements of cyst nuclei may provide the force necessary for karyogamy to occur and promote mixing of nuclear contents.

To determine how plasmid distribution changes relative to nuclear configuration, FISH was performed on encysting cultures to detect transfected plasmids (Fig. 2, H to L). As in Fig. 1, trophozoites contained plasmids in only one of the two nuclei (Fig. 2H). In 32% of cysts ($n = 50$), two nuclear pairs were present at opposite ends of the cell, with plasmids in one nucleus of each pair (Fig. 2I). In 21% of cysts, two nuclear pairs were present together at the same end of the cell, with plasmids in one nucleus of each pair (Fig. 2J). In both of these categories (Fig. 2, I and J), plasmids were detected in only one nucleus in each pair, indicating that the pairs are nondughters (derived from different parental nuclei). In 19% of cysts, three nuclei were present, presumably in the process of fusion, with plasmids in two of the three nuclei (Fig. 2K). Finally, in 27% of cysts, plasmids were seen in three of the four paired nuclei (Fig. 2L). Because all of these configurations appear at roughly the same time, the processes of nuclear migration, fusion, plasmid exchange, and fission may occur very quickly.

Cysts were immunolabeled with antibodies to α -tubulin to identify major microtubule arrays associated with cyst nuclei. Again, we saw three major classes of encysting cells: two pairs of nuclei far apart, two pairs close together, and

three nuclei (Fig. 3, A to C). Three-dimensional projections of these images (movies S4 to S6) demonstrate that the three nuclei seen in Fig. 3C are indeed three separate nuclei and not two nuclei superimposed on one another. The cyst microtubule cytoskeleton differs from the trophozoite cytoskeleton, because many of the complex structures seen in trophozoites are absent in the cyst (Fig. 3, A to C). Disassembled fragments of the ventral disk and internal axonemes are prominent (Fig. 3B). As shown by TEM (Fig. 1E) and immunofluorescence (Fig. 3, A to C), the nuclear pairs are associated with microtubule bundles, especially internal axonemes, throughout cyst formation, suggesting that the axonemes may have a distinctive function in *Giardia*.

On the basis of these data, we propose a model for plasmid transfer during encystation (Fig. 3D). First, at the induction of encystation, the nuclei divide with two separate spindles. This division generates a cyst with two pairs of non-daughter nuclei at opposite ends of the cell. Each pair of nuclei appears to be associated with a bundle of internal axonemes and moves together within the cyst (Figs. 1E and 2, B to G). One pair of nuclei then migrates to the opposite end of the cell to join the other nuclear pair. At some point, karyogamy between two non-daughter nuclei can occur, resulting in the appearance of three nuclei. This three-nuclear stage is commonly observed with all three nuclei in close proximity to one another, so the actual fusion probably occurs after the nuclei have migrated together at one end of the cell. Because we have never seen plasmids in all four nuclei of cysts, we conclude that fusion is typically restricted to only one set of nuclei. Alternatively, fusion may occur between both pairs of non-daughter nuclei, but may only occasionally result in genetic exchange. If fusion were occurring between daughter nuclei, it would not be detected as an increase in number of labeled nuclei in our FISH experiments. Finally, because fully mature cysts have four nuclei before excystation, we deduce that nuclear fission and the re-formation of four separate nuclei must follow karyogamy and nucleoplasm exchange.

For karyogamy to lead to genome homogenization, we assume that an event like homologous recombination occurs after the nuclear envelopes fuse. Thus, we predict that the expression of genes involved in DNA repair and/or homologous recombination should increase during encystation. Although meiosis has not been described in *Giardia*, analysis of its sequenced genome revealed homologs of meiosis-specific genes (HMGs), including Hop1, Spo11, Dmc1a, Dmc1b, and Mnd1 (5, 12). Nonetheless, it is possible that these genes had a nonmeiotic function in the last common ancestor of *Giardia* and other eukaryotes; indeed, many researchers believe that the eukaryotic meiosis machinery originally evolved from genes involved in DNA damage repair (13). The HMGs in *Giardia* may therefore facilitate the exchange of genomic

material within and between the cell's two nuclear genomes after karyogamy.

To assess the expression and localization of the protein products of HMGs, we constructed C-terminal green fluorescent protein (GFP) fusion constructs of these five genes under their native promoters. We observed DMC1A::GFP localization to the four cyst nuclei (fig. S1, A to C). However, no GFP was visible in the nuclei of the trophozoites (fig. S1, D to F). This same pattern was seen for SPO11::GFP (fig. S2, A and B) and HOP1::GFP (see below). In contrast, DMC1B::GFP (fig. S1, G to L) and MND1::GFP (fig. S2, C and D) localize to the nuclei of both trophozoites and cysts. Interestingly, although *Giardia* contains two putative homologs of the meiosis-specific RecA family member Dmc1, no homologs of Rad51, the somatic and meiotic

Fig. 4. The timing of HOP1::GFP expression during encystation. After 3 hours (A) of exposure to encystation conditions, no HOP1::GFP (green) is detected, but immunolabeling with antibodies to CWP (red) shows a small amount of CWP present in ESVs. After 12 hours of encystation (B), the amount of CWP detected in ESVs has increased, but no HOP1::GFP is visible. By 15 hours (C), HOP1::GFP is detected and colocalizes to two DAPI (blue)-stained nuclei, while the cell is filled with vesicles containing CWP. When encystation is complete at 24 hours (D), HOP1::GFP is seen in the four-cyst nuclei, and the cyst wall is smooth and contiguous around the periphery of the cell.

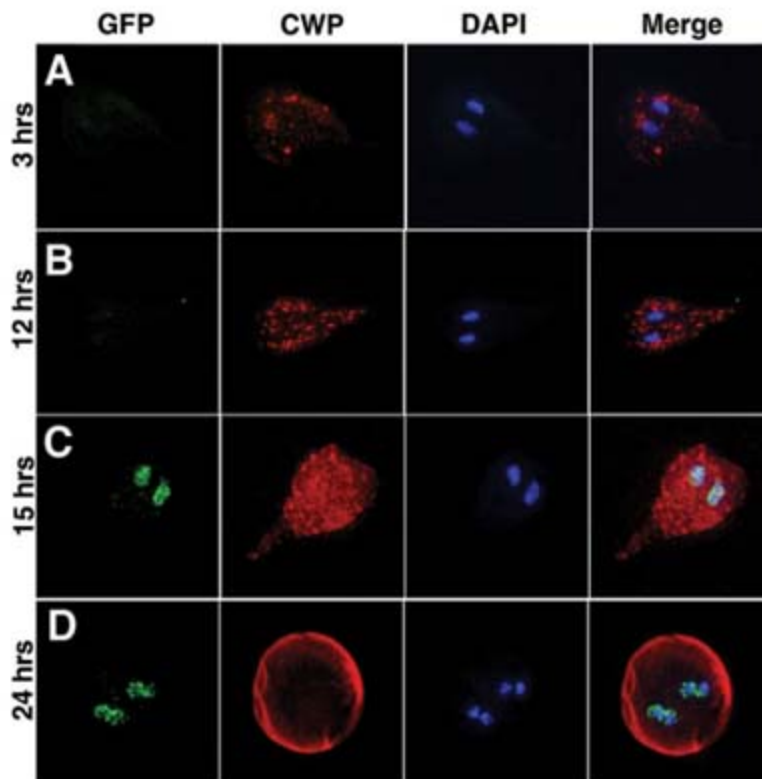
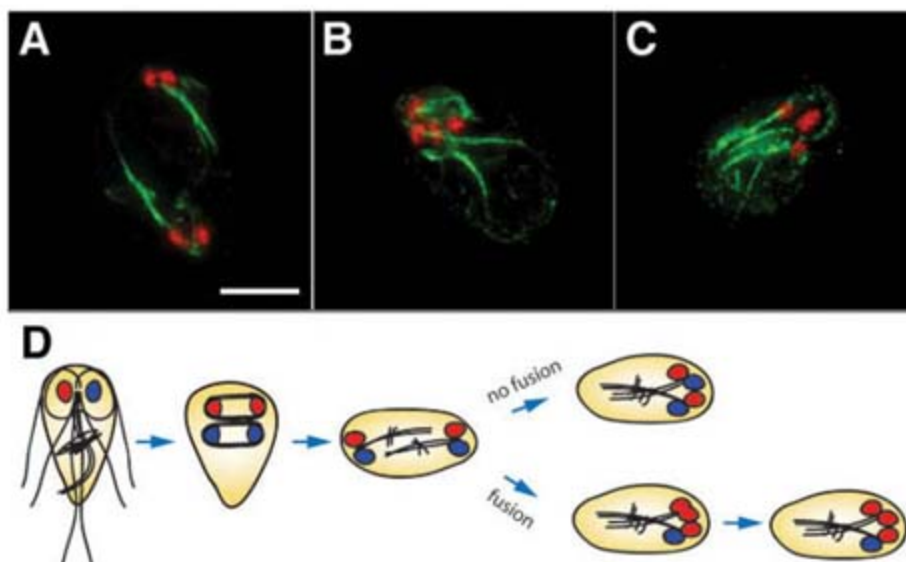


Fig. 3. The microtubule cytoskeleton in cysts and a model of nuclear movement. (A to C) Encysting cells were immunolabeled with antibodies to α -tubulin (green) and the DNA stained with DAPI (false color, red). (A) Two pairs of nuclei are at opposite ends of the cell. (B) Four nuclei clustered together at one end of the cell. (C) Only three nuclei are distinguishable in the cluster. Bar, 5 μ m. Also see movies S4 to S6. (D) A model for plasmid exchange during encystation. A trophozoite has two independently inherited nuclei, blue or red, and only one (red) contains plasmids. Each nucleus divides during encystation, resulting in two pairs of non-daughter nuclei at opposite ends of the cell. After nuclear migration, fusion occurs between non-daughter nuclei (i.e., between a red and a blue nucleus), leading to transfer of genetic material. Subsequently, fused nuclei undergo fission and move apart, resulting in three (red) nuclei containing plasmids.



RecA family member, have been identified (12). Because DMC1B is expressed in vegetatively growing trophozoites and cysts whereas DMC1A is cyst-specific, we postulate that DMC1B may have a Rad51-like function and be involved in somatic DNA damage repair. Similarly, we suspect that although Mnd1 is involved in meiosis-specific recombination in other organisms, it may have a more general role in *Giardia*.

To determine when HOP1::GFP begins to be localized to the cyst nuclei, we monitored encystation over 24 hours with an antibody to cyst wall protein (CWP), which is transported via encystation-specific vesicles (ESVs) to the exterior of the cell and deposited to form the cyst wall (14). Three hours after induction of encystation, CWP was detected in ESVs in the cytoplasm (Fig. 4A), and the number of ESVs continued to increase through the first 12 hours (Fig. 4B). After 15 hours, cells were filled with ESVs, and HOP1::GFP was first detected in the two nuclei of the encysting trophozoite (Fig. 4C). By 24 hours (Fig. 4D), mature cysts with four nuclei and a defined wall were present, and HOP1::GFP was localized to all four nuclei. Thus, HOP1, which is involved in binding double-strand breaks during meiosis in yeast (15), is first detected in the nuclei in *Giardia* during encystation and persists indefinitely in cysts.

Karyogamy during encystation, if accompanied by homologous recombination and/or gene conversion driven by the HMGs, could provide a mechanism by which *Giardia* maintains low levels of allelic heterozygosity. This parasexual process, which we call diplomixis, appears to be unique to *Giardia*, although we predict that it occurs in other members of the order Diplomonadida. Unlike automixis, diplomixis is not accompanied by meiotic genome reduction and the subsequent fusion of gametes from the same parent, as is found in the sexual or parasexual life cycle of other organisms (16). It is also possible that rare meiotic events occur in the wild, as suggested by a recent *Giardia* population study (17). Understanding the functions of the HMGs and the behavior of chromosomes as nuclei fuse will be essential to test for the occurrence of homologous recombination. A deeper understanding of the roles of these genes, as well as others involved in karyogamy, may shed light on the evolution of meiosis and provide new targets for drug treatments.

References and Notes

1. L. Savioli, H. Smith, A. Thompson, *Trends Parasitol.* **22**, 203 (2006).
2. R. D. Adam, T. E. Nash, T. E. Wellems, *Nucleic Acids Res.* **16**, 4555 (1988).
3. K. S. Kabnick, D. A. Peattie, *J. Cell Sci.* **95**, 353 (1990).

4. L. Z. Yu, C. W. Birky, R. D. Adam, *Eukaryot. Cell* **1**, 191 (2002).
5. H. G. Morrison *et al.*, *Science* **317**, 1921 (2007).
6. S. Teodorovic, J. M. Braverman, H. G. Elmdorf, *Eukaryot. Cell* **6**, 1421 (2007).
7. R. D. Adam, *Clin. Microbiol. Rev.* **14**, 447 (2001).
8. R. Bernander, J. E. Palm, S. G. Svärd, *Cell. Microbiol.* **3**, 55 (2001).
9. M. S. Sagolla, S. C. Dawson, J. J. Mancuso, W. Z. Cande, *J. Cell Sci.* **119**, 4889 (2006).
10. C. W. Birky, *Genetics* **144**, 427 (1996).
11. Materials and methods are available as supporting material on Science Online.
12. M. A. Ramesh, S. B. Malik, J. M. Logsdon, *Curr. Biol.* **15**, 185 (2005).
13. A. M. Villeneuve, K. J. Hillers, *Cell* **106**, 647 (2001).
14. D. S. Reiner, H. Douglas, F. D. Gillin, *Infect. Immun.* **57**, 963 (1989).
15. N. M. Hollingsworth, L. Goetsch, B. Byers, *Cell* **61**, 73 (1990).
16. A. S. Kondrashov, *Annu. Rev. Ecol. Syst.* **28**, 391 (1997).
17. M. A. Cooper, R. D. Adam, M. Worobey, C. R. Sterling, *Curr. Biol.* **17**, 1984 (2007).
18. We thank the Cande and Dawson labs for discussion, as well as E. Slawson for help with experiments. We gratefully acknowledge funding from the NIH (grants A1054693 to W.Z.C. and 1F32GM078971 to M.K.P.) and the NSF (predoctoral fellowship to M.L.C.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/319/5/869/1530/DC1

Materials and Methods

Figs. S1 and S2

References

Movies S1 to S6

4 December 2007; accepted 11 February 2008

10.1126/science.1153752

Direct Visualization of Horizontal Gene Transfer

Ana Babić,^{1,2*} Ariel B. Lindner,^{1,2} Marin Vulić,^{1,2†} Eric J. Stewart,^{1,2†} Miroslav Radman^{1,2,3‡}

Conjugation allows bacteria to acquire genes for antibiotic resistance, novel virulence attributes, and alternative metabolic pathways. Using a fluorescent protein fusion, SeqA-YFP, we have visualized this process in real time and in single cells of *Escherichia coli*. We found that the F pilus mediates DNA transfer at considerable cell-to-cell distances. Integration of transferred DNA by recombination occurred in up to 96% of recipients; in the remaining cells, the transferred DNA was fully degraded by the RecBCD helicase/nuclease. The acquired integrated DNA was tracked through successive replication rounds and was found to occasionally split and segregate with different chromosomes, leading to the inheritance of different gene clusters within the cell lineage. The incidence of DNA splitting corresponds to about one crossover per cell generation.

Together with transformation and phage-mediated transduction, conjugation is a key mechanism for horizontal gene transfer in bacteria (1). The first evidence for sex by conjugation in *E. coli* was provided by Lederberg, who obtained prototrophic progeny by mixing two different auxotrophic parents (2). Since then, the phenomenon of horizontal gene transfer has been shown to be responsible for widespread transfer among bacterial populations of genes conferring antibiotic resistance, metabolic functions, and virulence determinants.

Conjugational DNA transfer is driven by the F plasmid unidirectionally from an F⁺ donor cell to an F⁻ recipient cell. The F plasmid contains all the genes required for conjugation (e.g., medi-

ating the contact between donor and recipient cells) and for regulation of DNA mobilization and its unidirectional transfer (3). At low frequencies, the F plasmid can integrate into the chromosome of the host cell, giving rise to an Hfr (high frequency of recombination) strain (4). Chromosomal genes of the Hfr bacterium can be mobilized and transferred to a recipient. In some cases, F can excise from the chromosome of Hfr, creating an F['] molecule that carries chromosomal genes as well as the conjugation genes (5). Both Hfr and F['] can serve as DNA vehicles in horizontal gene transfer between bacteria.

The contact between mating cells is mediated by a tube-like structure known as the F pilus (3). DNA is transferred from the donor to the recip-

ient in single-stranded form and converted to duplex DNA by the synthesis of the complementary strand in the recipient cell. Once the conjugational transfer ceases, double-stranded donor DNA is either circularized (in the case of F['] transfer) or, in the case of Hfr transfer, incorporated into the recipient chromosome by RecA-dependent homologous recombination or degraded by RecBCD exonuclease (3, 6).

Many aspects of the mechanism and consequences of conjugation remain unresolved, including the role of the F pilus in DNA transfer during conjugation, the fate of the transferred DNA, the global frequency of the horizontal gene transfer (versus the frequency of inheritance of individual genetic markers), and the pattern of inheritance of donor DNA present in the initial transconjugant cell. To address these questions, we have developed an experimental system that enables us to distinguish the transferred donor DNA from both donor and recipient DNA, and to visualize DNA transfer and recombination by means of fluorescence microscopy in real time,

¹INSERM U571, Paris F-75015, France. ²Université Paris Descartes Faculté de Médecine, Paris F-75015, France. ³Mediterranean Institute for Life Sciences, Meštrovićevo Šetalište bb, 21000 Split, Croatia.

*Present address: Unité Plasticité du Génome Bactérien, CNRS URA 2171, Institut Pasteur, 25 rue de Dr. Roux, Paris 75724, France.

†Present address: Department of Biology, Northeastern University, 309 Mugar Hall, 360 Huntington Avenue, Boston, MA 02115, USA.

‡To whom correspondence should be addressed. E-mail: radman@necker.fr

at the level of individual living cells. This tool allowed us to quantify the ongoing transfer of DNA during conjugation and to acquire time-lapse movies that follow the fate of the newly acquired DNA in individual cells through any number of cell divisions.

SeqA is a negative regulator of initiation of *E. coli* replication, with a high affinity for DNA that has been hemimethylated (i.e., only one strand in the duplex DNA is methylated) by Dam methylase at 5' GATC sequences (7–9). We mated Dam methylation-proficient donor (Hfr or F') cells with methylation-deficient recipient cells, producing the SeqA-YFP (yellow fluorescent protein) fusion protein, which enabled us to specifically and permanently label only the transferred DNA. Once methylated donor DNA is transferred to the Dam⁻ recipient as a single strand and converted to a DNA duplex by the synthesis of the complementary nonmethylated strand, the acquired DNA is expected to remain permanently hemimethylated and therefore bound by SeqA-YFP. Indeed, we have previously shown the binding of SeqA to conjugally transferred DNA by immunofluorescence on fixed cells (10).

The SeqA-YFP translational fusion was constructed and placed in the chromosome under the control of the native *seqA* promoter, replacing the resident *seqA* gene (11). As expected, in Dam⁺ cells SeqA-YFP localizes to DNA replication forks to form compact foci (Fig. 1A). As DNA methylation lags considerably behind DNA synthesis, the newly replicated DNA is transiently hemimethylated DNA and therefore transiently bound by SeqA-YFP (Fig. 1A) (12–14). In contrast, in Dam⁻ cells SeqA-YFP fluorescence is diffuse (Fig. 1B). The number of SeqA-YFP foci in Dam⁺ cells varied from one to two in minimal medium, and up to eight in LB (Luria Bertani) medium, as expected from the occurrence of multiple replication forks in rich medium (13, 15). Foci were positioned at the half- or quarter-length positions in the cell. Both the number and position of SeqA-YFP foci corresponded to those previously reported for SeqA (13, 14), which suggests that SeqA-YFP is functional in terms of DNA binding specificity.

To investigate the localization of SeqA-YFP protein during conjugation, we placed a mixture of dam⁺ donors and dam⁻ seqA-yfp recipients on a nutrient-containing agarose slab in a sealed-cavity microscope slide and observed it with time-lapse fluorescence microscopy. About 5 min after mixing the parental cells, distinct SeqA-YFP foci started to appear in recipient cells (Fig. 1, D and E). During the next 30 to 40 min, almost all recipients in the vicinity of donors acquired intense fluorescent foci of the SeqA-YFP due to the sequestration of previously diffused SeqA-YFP by the transferred DNA (Fig. 1E and movie S1). A single recipient cell could receive DNA more than once, as shown by the presence of independent, well-separated foci arising at different time points at different positions in the cell (fig. S1 and movie S1).

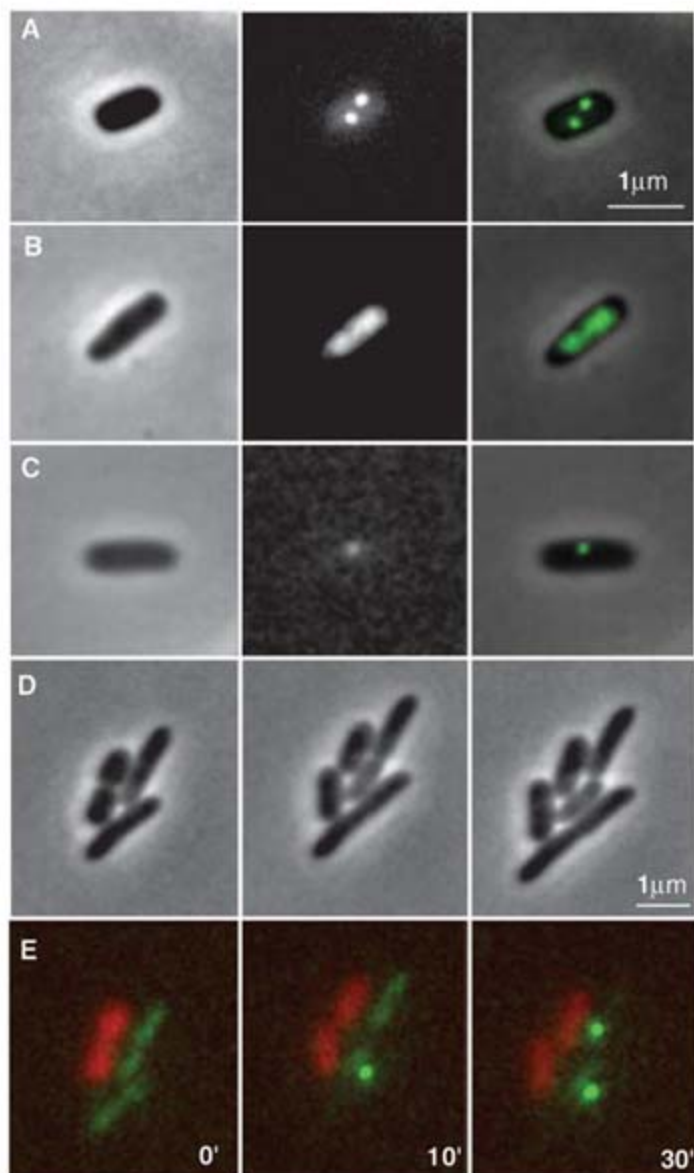
As expected, when dam⁻ donors were used in conjugation, no SeqA-YFP foci were observed in the dam⁻ recipients. Also, no SeqA-YFP foci were seen when donors lacking the structural pilus protein (TraA) were used and were therefore transfer-deficient (16). These results confirm that SeqA-YFP foci formation reflects the entry of a methylated donor DNA strand into the dam⁻ recipient during conjugation. Although the conjugal transfer of the extrachromosomal R plasmid was previously observed by fluorescence of a specific locus on the plasmid (17), our system allows visualization of transfer of any extrachromosomal or chromosomal DNA in a sequence-independent fashion.

To visualize the expression of conjugally transferred DNA in the recipient cells, we constructed the ANB35 donor strain of *E. coli*, which carries the tetracycline (Tet) promoter-driven monomeric red fluorescent protein (*mrfp1*) gene (18) and *tetR* (repressor) gene (11). In the absence of induction, donor cells exhibit weak red fluorescence because of the leakiness of the Tet promoter (movie S2). When ANB35 donors were mated with dam⁻ seqA-yfp recipients, we first observed the appearance of the SeqA-YFP

focus, denoting DNA transfer, followed by the red fluorescence in the recipient cell, showing *mrfp1* gene expression, about 2 hours after the transfer of the *mrfp1* gene. This expression of the *mrfp1* gene was detected in 25 of 307 mated recipients (8.1%). Not all successful transfers would have included the *mrfp1* gene, because conjugational transfer ceases at random points, and the probability of transferring a specific marker decreases exponentially with its distance from origin of transfer *oriT* (19). The recipient cell eventually became intensely red relative to the fluorescence of the donor (movie S2). Upon entry of the donor DNA into the recipient cell lacking the TetR repressor, the transcription of *mrfp1* occurred before sufficient TetR repressor was expressed and functional. The recipient cell therefore became intensely red before repression started. This phenomenon is known as induction of gene expression by conjugation, or zygotic induction (20).

The F pilus is a tubular organelle, extruding from the surface of the donor cell, that is required for conjugation (21). Some authors have argued that the F pilus serves only to bring the mating cells in direct contact, and that DNA is trans-

Fig. 1. Localization of SeqA-YFP in DNA replication and in conjugation. (A) Dam-proficient cells. (B) Dam-deficient cells. (C) Recipient with SeqA-YFP focus after conjugation with F donor. Left, phase contrast image; center, fluorescence image; right, overlay between the phase contrast image and the fluorescence image represented in green. (D and E) Real-time conjugation [(D), phase contrast; (E), fluorescence overlay] of donors (red cells) with recipient (green cells) as followed by time-lapse microscopy at 0, 10, and 30 min after plating on nutrient-agarose cavity slide.



ferred subsequently through a pore in the membrane (22, 23). Others have shown that DNA can be transferred from the donor to the recipient even when mating cells were apparently physically separated (24). In three experiments, we observed SeqA-YFP foci appearing in 43 of 753 recipients that were not in direct cell wall contact; the maximal observed distance to the closest donor was 12 μm (Fig. 2 and movie S3). No distant transfer of DNA was observed in the presence of 0.01% SDS (which depolymerizes F pili) or when using the transfer-deficient *traA*⁻ donors (which lack the pili). Clearly, the appearance of distant SeqA-YFP foci depends on conjugation and on the presence of F pili. Thus, our results offer evidence supporting the idea that the F pilus, in addition to establishing the contact between mating cells, serves as a channel for single-stranded DNA transfer during conjugation.

In conjugations with recombination-deficient recipients (*recA*⁻), transferred DNA is known to be degraded by the RecBCD exonuclease (6, 25). RecBCD is a highly processive enzyme degrading single- and double-stranded linear DNA (26, 27). The kinetics of degradation of conjugally transferred DNA have so far only been measured in large cell populations (6, 25) and not at the level of single cells.

In our experiments we found that the SeqA-YFP focus fluorescence intensity in the recipient is proportional to the time of active conjugation, that is, to the length of transferred DNA (fig. S2). The gradual disappearance of the fluorescence of

SeqA-YFP foci that we observed in recombination-deficient cells (movie S4) probably reflected DNA degradation mediated by the RecBCD helicase/nuclease complex. Indeed, Hfr DNA transferred into *recA*⁻*recD*⁻ recipients was largely left unchanged, and no degradation was seen in 459 of 471 recipients [i.e., $97 \pm 0.24\%$ (SEM) in three experiments] 4 hours after interruption of mating. By contrast, in *recA*⁻ recipients, transferred DNA was rapidly broken down in 313 of 343 recipients (i.e., $91.3 \pm 4.1\%$ in three experiments). DNA degradation over time in most cases was linear, yet individual cells exhibited variable degradation rates (fig. S3A). Occasionally (for 9 of 135 foci in two experiments) we saw a delay in the onset of degradation (fig. S3B). Interestingly, a subpopulation of SeqA-YFP foci ($8.7 \pm 4.1\%$ for 343 tracked foci) exhibited an initial decrease in fluorescence due to degradation of DNA but eventually reached a stable fluorescence level, even 4 hours after the interruption of mating (fig. S3C). This contrasts with the findings that no genuine genetic recombinants can be obtained in crosses with *recA*⁻ recipients (28). In these cases, the nonreplicative Hfr DNA somehow must be end-protected—for example, by circularization via microhomologies, palindromic end structures, or protein binding.

The fate of transferred DNA in recombination-proficient (*recA*⁺) recipients was followed after interruption of mating (11). In *E. coli*, all free linear DNA is rapidly degraded by the RecBCD exonuclease (27); therefore, SeqA-YFP foci that remain stable in *recA*⁺ recipients for at least 4

hours were considered as recombined DNA. Transferred DNA followed two fates in the *recA*⁺ recipients: In 11 of 336 mated recipients ($3.3 \pm 0.83\%$ in three experiments), SeqA-YFP foci disappeared, indicating that the donor DNA was degraded (fig. S4). However, in the majority ($96.7 \pm 0.83\%$), the fluorescence intensity remained constant or slowly increased for the higher-intensity foci. This is coherent with the binding of newly synthesized SeqA-YFP to available GATC sequences of large conjugated fragments, whereas small fragments are readily saturated by available SeqA-YFP (fig. S4).

This direct, physical, measure of recombination shows substantially higher recombination frequencies than those measured genetically as the acquisition of specific genetic markers (ranging from 10 to 30% depending on specific genetic marker, donor strain, and mating conditions) (29, 30). Our results show that conjugational recombination is an extremely efficient process when donors and recipients are essentially genetically identical strains.

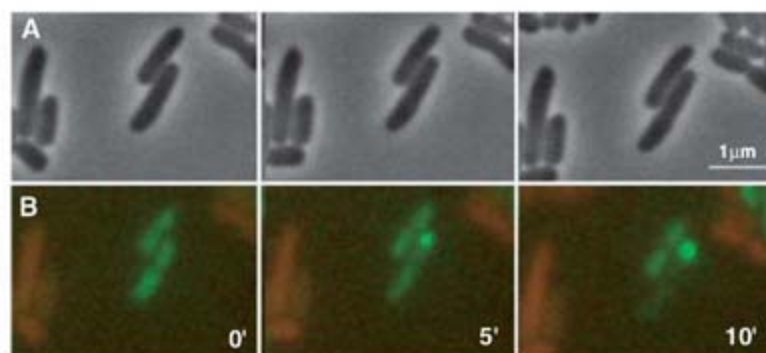
Remarkably, SeqA-YFP foci frequently split within single cell lineages (Fig. 3 and movie S5). The sum of fluorescence intensities of newly formed foci was not statistically different from the intensity of the ancestral focus (*t* test, $P = 0.62$, $n = 20$); that is, the splitting events were conservative, showing no evidence of significant DNA degradation (Fig. 3). Most likely, the splitting of SeqA-YFP foci reflects the physical breakage of the initial transferred DNA strand into two DNA fragments (11). This is reminiscent of reciprocal crossovers (sister chromatid exchange) because the splitting frequency is proportional to the DNA size and occurs only in recombination-proficient cells (11). Extrapolation of the splitting frequency of the acquired DNA to the entire genome suggests that there are about 1.4 crossovers (sister or cousin molecule exchange) per cell generation.

This work allowed us to visualize and quantify the DNA of any sequence as it is being transferred from one individual cell to another, and to watch its stable genomic acquisition via genetic recombination (horizontal gene transfer) in real time. The observations also implicated F pilus-mediated conjugation at cell-to-cell distances of up to 12 μm and permitted estimation of the frequency of intragenomic crossover events, or sister-chromatid exchanges. This experimental system can now be applied to monitor horizontal gene transfer by indefinitely following the fate of DNA acquired in intra- and interspecies crosses.

References and Notes

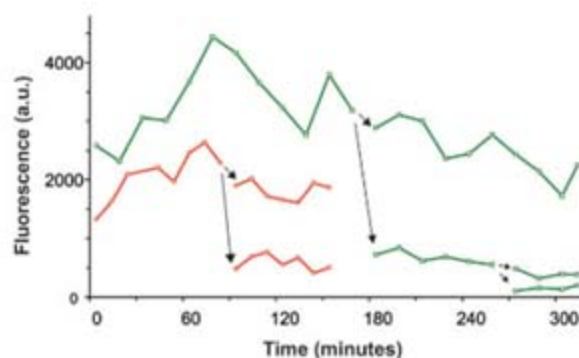
- H. Ochman, J. G. Lawrence, E. A. Groisman, *Nature* **405**, 299 (2000).
- J. Lederberg, E. L. Tatum, *Cold Spring Harb. Symp. Quant. Biol.* **11**, 113 (1946).
- N. Willetts, R. Skurray, in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, F. C. Neidhardt, Ed. (American Society for Microbiology, Washington, DC, 1987), pp. 1110–1113.

Fig. 2. Representative example of conjugational DNA transfer without visible cell contact between donors and recipients. Recipients and donors were plated on the nutrient-agarose slab without previously being in contact and were observed by time-lapse fluorescence microscopy. (A) Phase contrast



images. (B) Overlay of fluorescence images of recipients (green cells) and donors (red cells) at 0, 5, and 10 min after plating on LB-agarose in a cavity slide.

Fig. 3. Two representative cases of the SeqA-YFP focus splitting, differentiated by color. SeqA-YFP foci were acquired during 30 min of mating on a nitrocellulose filter with a recombination-proficient recipient. Time 0 represents the interruption of mating by nalidixic acid and transfer under the microscope for observation. Times of splitting are indicated by arrows; a.u., arbitrary units. Fluorescence 0 is the average background SeqA-YFP fluorescence level of nonconjugating recipients. Each point represents an integral (total) fluorescence of SeqA-YFP focus.



4. W. Hayes, *Cold Spring Harb. Symp. Quant. Biol.* **18**, 75 (1953).
5. K. B. Low, *Methods Enzymol.* **204**, 43 (1991).
6. R. G. Lloyd, C. Buckman, *Genetics* **139**, 1123 (1995).
7. T. Brendler, A. Abeles, S. Austin, *EMBO J.* **14**, 4083 (1995).
8. M. Lu, J. L. Campbell, E. Boye, N. Kleckner, *Cell* **77**, 413 (1994).
9. S. Slater *et al.*, *Cell* **82**, 927 (1995).
10. M. Kohiyama, S. Hiraga, I. Matic, M. Radman, *Science* **301**, 802 (2003).
11. See supporting material on Science Online.
12. J. L. Campbell, N. Kleckner, *Cell* **62**, 967 (1990).
13. S. Hiraga, C. Ichinose, H. Niki, M. Yamazoe, *Mol. Cell* **1**, 381 (1998).
14. T. Onogi, H. Niki, M. Yamazoe, S. Hiraga, *Mol. Microbiol.* **31**, 1775 (1999).
15. C. Helmstetter, S. Cooper, O. Pierucci, E. Revelas, *Cold Spring Harb. Symp. Quant. Biol.* **33**, 809 (1968).
16. L. S. Frost, W. Paranchych, N. S. Willetts, *J. Bacteriol.* **160**, 395 (1984).
17. T. D. Lawley, G. S. Gordon, A. Wright, D. E. Taylor, *Mol. Microbiol.* **44**, 947 (2002).
18. R. E. Campbell *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 7877 (2002).
19. G. R. Smith, *Cell* **64**, 19 (1991).
20. F. Jacob, E. L. Wollman, *Ann. Inst. Pasteur* **91**, 486 (1956).
21. J. Manchak, K. G. Anthony, L. S. Frost, *Mol. Microbiol.* **43**, 195 (2002).
22. M. Achtman, G. Morelli, S. Schwuchow, *J. Bacteriol.* **135**, 1053 (1978).
23. M. M. Panicker, E. G. Minkley Jr., *J. Bacteriol.* **162**, 584 (1985).
24. L. C. Harrington, A. C. Rogerson, *J. Bacteriol.* **172**, 7263 (1990).
25. S. Delmas, I. Matic, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 4564 (2006).
26. A. K. Eggleston, S. C. West, *Curr. Biol.* **7**, R745 (1997).
27. S. C. Kowalczykowski, D. A. Dixon, A. K. Eggleston, S. D. Lauder, W. M. Rehrauer, *Microbiol. Rev.* **58**, 401 (1994).
28. B. Low, *Proc. Natl. Acad. Sci. U.S.A.* **60**, 160 (1968).
29. L. S. Baron, P. Gemski, E. M. Johnson, J. A. Wohlhieter, *Bacteriol. Rev.* **32**, 362 (1968).
30. C. Rayssiguier, C. Dohet, M. Radman, *Biochimie* **73**, 371 (1991).
31. Supported by fellowships from ARC and Chancellerie des Universites de Paris and a fellowship supplement from the Croatian Ministry of Science, Education and Sports (A.B.); postdoctoral fellowships from EMBO and Marie Curie program (A.B.L.); postdoctoral fellowships from Ligue Contre le Cancer and Inserm (Poste Vert) (M.V.); postdoctoral fellowships from EMBO, Fondation pour la Recherche Medicale, and Inserm (E.J.S.); and Inserm and Institut Necker (contract Mixis-Pliva).

Supporting Online Material

www.sciencemag.org/cgi/content/full/319/5869/1533/DC1

Materials and Methods

SOM Text

Figs. S1 to S4

Movies S1 to S5

References

28 November 2007; accepted 5 February 2008

10.1126/science.1153498

Neurokinin 1 Receptor Antagonism as a Possible Therapy for Alcoholism

David T. George,^{1*} Jodi Gilman,^{1*} Jacqueline Hersh,^{1*} Annika Thorsell,^{1*} David Herion,¹ Christopher Geyer,² Xiaomei Peng,³ William Kielbasa,³ Robert Rawlings,¹ John E. Brandt,³ Donald R. Gehlert,³ Johannes T. Tauscher,³ Stephen P. Hunt,⁴ Daniel Hommer,¹ Markus Heilig^{1†}

Alcohol dependence is a major public health challenge in need of new treatments. As alcoholism evolves, stress systems in the brain play an increasing role in motivating continued alcohol use and relapse. We investigated the role of the neurokinin 1 receptor (NK1R), a mediator of behavioral stress responses, in alcohol dependence and treatment. In preclinical studies, mice genetically deficient in NK1R showed a marked decrease in voluntary alcohol consumption and had an increased sensitivity to the sedative effects of alcohol. In a randomized controlled experimental study, we treated recently detoxified alcoholic inpatients with an NK1R antagonist (LY686017; $n = 25$) or placebo ($n = 25$). LY686017 suppressed spontaneous alcohol cravings, improved overall well-being, blunted cravings induced by a challenge procedure, and attenuated concomitant cortisol responses. Brain functional magnetic resonance imaging responses to affective stimuli likewise suggested beneficial LY686017 effects. Thus, as assessed by these surrogate markers of efficacy, NK1R antagonism warrants further investigation as a treatment in alcoholism.

Alcohol use accounts for 4% of global disease burden (1). Alcohol dependence, or alcoholism, is characterized by a chronic relapsing course, in which alcohol-associated cues and stress are known relapse triggers (2–6). Recent research suggests that neural systems mediating behavioral stress responses may offer useful targets for pharmacotherapy of alcohol-

ism. In animal models, excessive alcohol consumption that results from a history of alcohol dependence is accompanied by increased behavioral sensitivity to stress (7). Up-regulated corticotropin-releasing hormone (CRH) signaling in extrahypothalamic brain sites contributes to these dependence-induced changes, but other stress-related neurotransmitters may also play a role.

One such neurotransmitter is substance P (SP), which together with its preferred neurokinin 1 receptor (NK1R) is highly expressed in brain areas involved in stress responses and drug reward, including the hypothalamus, amygdala, and nucleus accumbens. In rodents, psychological stressors induce release of SP in the amygdala, whereas genetic deletion or pharmacological blockade of NK1R inhibits the associated behavioral responses (8). Furthermore,

genetic deletion of NK1Rs causes a loss of conditioned place preference for opiates and opiate self-administration (9, 10). In humans, the NK1 antagonist GR205171 reduces symptoms of social anxiety and suppresses brain responses to the Trier Social Stress Test (TSST) (11). Together, these findings suggest that blockade of NK1Rs might modulate stress- and reward-related processes of importance for excessive alcohol use and relapse. To our knowledge, no data are presently available to address this hypothesis.

We first explored preclinically whether inactivation of NK1R might modulate stress- and reward-related processes that impact alcohol use. We chose a genetic inactivation strategy, because available NK1R antagonists have limited activity in rats and mice, because of insufficient NK1R amino acid homology between humans and these rodent species (8). We evaluated NK1R null-mutant mice for voluntary alcohol consumption, alcohol sensitivity, and alcohol metabolism (12). NK1R null mice (13) were back-crossed into a C57BL/6 background for 10 generations to ensure that there was adequate voluntary alcohol consumption in control animals (14). We used a two-bottle free-choice model with increasing alcohol concentration, and alcohol was continuously available. Wild-type littermates (++) ultimately consumed in excess of 10 g alcohol/kg of body weight per day at the end of an escalation procedure in which alcohol concentration was gradually increased from 3 to 15% over 60 days.

Alcohol consumption by NK1R^{-/-} mice was markedly lower than that by wild-type controls (Fig. 1A). The difference was most prominent at higher alcohol concentrations, at which consumption motivated by pharmacological alcohol effects dominates over intake for taste, calories, or other nonpharmacological effects (14). Alcohol consumption by heterozygous (+/-) mice was similar to wild-type controls, highlighting the necessity for near-complete inactivation of NK1Rs to suppress alcohol consumption. Sev-

¹Laboratory of Clinical and Translational Studies, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Bethesda, MD 20892, USA. ²Department of Nursing, Mark O. Hatfield Clinical Research Center, National Institutes of Health, Bethesda, MD 20892, USA. ³Lilly Research Laboratories, Indianapolis, IN 46285, USA. ⁴Department of Anatomy and Developmental Biology, University College London, London WC1E 6BT, UK.

*These authors contributed equally.

†To whom correspondence should be addressed. E-mail: markus.heilig@mail.nih.gov

eral observations indicated that the reduction in alcohol intake was specific. Thus, relative preference for alcohol was also markedly reduced [$F_{2,45} = 13.6$, $P < 0.0001$], whereas water intake was unaffected by genotype ($F_{2,45} = 0.54$, $P = 0.71$). Additional control experiments indicated that the reduction in alcohol intake was not caused by altered thirst or taste preference (12).

Alcohol sensitivity, measured as the time required to regain the righting reflex after a high dose of alcohol (3.5 g/kg), was markedly increased in $NK1^{-/-}$ mice (Fig. 1B). This profile is consistent with an abundance of animal data showing an inverse relation between alcohol sensitivity and motivation to consume alcohol, as well as human findings of an inverse relation between alcohol sensitivity and alcoholism risk (15).

To assess the potential clinical relevance of these results, we conducted an experimental study in human alcohol-dependent patients (fig. S1 and table S1), in which we evaluated the effects of NK1R antagonism on processes related to

relapse. We studied LY686017 (16), a high-affinity, selective NK1R antagonist that is orally available and brain penetrant (Fig. 2A). Preclinical pharmacology, safety, and human pharmacokinetics of LY686017 will be reported separately. Activity of LY686017 to reduce drinking was not assessed in preclinical experiments because, similar to other human NK1R antagonists, LY686017 has insufficient affinity for the mouse or rat NK1R. LY686017's brain penetrance and NK1R occupancy were established in a human positron emission tomography (PET) study of eight healthy volunteers (12). On the basis of these results, we used a dose of 50 mg daily, which yields a >90% blockade of central NK1R. LY686017 at this dose was well tolerated (table S2) (12), in agreement with prior reports indicating that NK1 antagonists as a class are safe and well tolerated (8).

Given the role of SP and NK1 receptors in stress and anxiety responses, we targeted subjects with high trait anxiety. Participants (25 per arm;

age: 21 to 65 years) had a diagnosis of alcohol dependence, alcohol problems as the primary complaint, alcohol use within the last month, and >39 on the Spielberger Trait Anxiety Inventory (STAI) (17). They were hospitalized throughout the study and had completed withdrawal treatment before entering the study if needed. For details, see (12).

We found that LY686017 suppressed spontaneous alcohol cravings, as measured by the Alcohol Urge Questionnaire (AUQ) (Fig. 2B). As expected, these cravings declined over time in the protected inpatient environment and were minimal in the majority of patients by the end of the 4-week study period. However, overall, the cravings were sufficiently intense to allow detection of the medication effect. Weekly ratings by a blinded observer likewise suggested that LY686017 had a beneficial effect on global improvement and severity measured by the Clinician's Global Impression (CGI) scale (Fig. 2C). In contrast, twice-weekly ratings collected to obtain measures of general psychopathology showed no treatment effects on anxious or depressive psychopathology ($F_{1,102} = 0.7$, $P = 0.40$). This suggests that the improvements observed might be specific for brain processes related to alcoholism.

We further measured craving responses to a combined stress (TSST) and alcohol-cue challenge (12) and found that treatment with LY686017 reduced the resulting AUQ craving response (Fig. 3A). Treatment with LY686017 also suppressed the concomitant cortisol response to the challenge (Fig. 3B). A study using PET has previously shown that GR205171, another NK1 antagonist, suppresses amygdala activation in response to the TSST in social phobics, but in that study, NK1R antagonism failed to produce effects on subjective, self-reported measures in response to the challenge (11). This shows that subjective responses to a challenge are transient, and their detection is critically dependent on the time point and the assessment instrument chosen. The fact that we detected consistent effects of LY686017 on both subjective and neuroendocrine responses to the craving challenge, therefore, supports the robustness of the LY686017 effect.

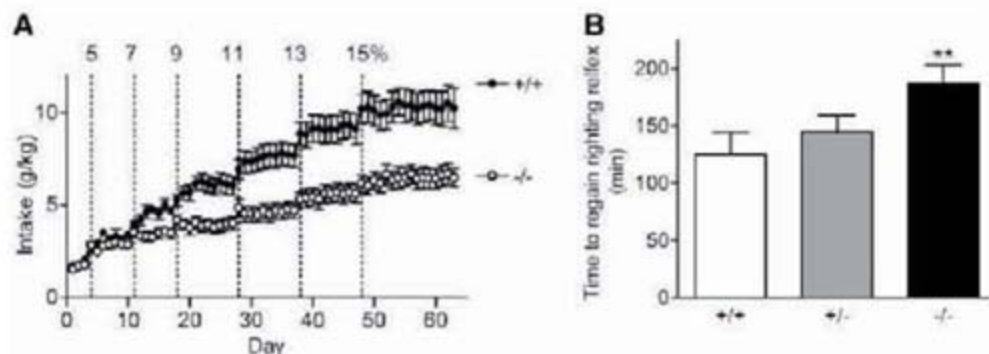
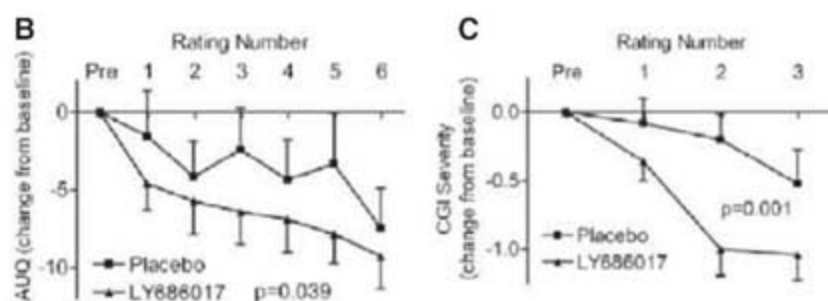
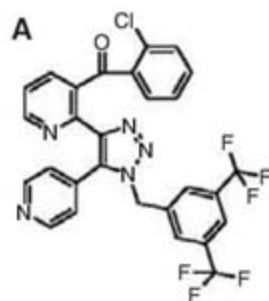


Fig. 1. Voluntary alcohol intake and alcohol sensitivity in NK1 null ($-/-$), heterozygous ($+/-$) and wild-type ($+/+$) mice. **(A)** Voluntary consumption in a two-bottle free-choice paradigm with continuous alcohol access. Alcohol was introduced as a 3% solution in water, and concentration was escalated over time, as indicated above the graph, to avoid taste aversion and to achieve pharmacologically active levels of alcohol consumption. The other bottle contained water. Data are means \pm SEM; $n = 16$ per group. The $+/-$ group is omitted for clarity; its alcohol consumption was virtually identical to that of $+/+$ mice. There was a main genotype effect ($F_{2,40} = 929.6$, $P < 0.00001$). On Tukey's post hoc test, $-/-$ mice differed from both $+/+$ and $+/-$ mice (both: $P < 0.001$); the latter two genotypes were virtually identical ($P = 0.92$). **(B)** Alcohol sensitivity in mice, measured as the time to regain the righting reflex after a 3.5-g/kg dose of alcohol. Data are means \pm SEM; $n = 12$ to 17/group. Time to regain the righting reflex was significantly longer for the $NK1^{-/-}$ mice as compared with their wild-type controls ($F_{2,42} = 7.8$, $P = 0.0014$; Tukey's post hoc $P = 0.001$ $-/-$ vs. $+/+$).

Fig. 2. Effect of the NK1R antagonist LY686017 on spontaneous cravings and global well-being in hospitalized alcoholics. **(A)** LY686017 (2-chloro-phenyl)-{2-[5-pyridin-4-yl-1-(3,5-bis(trifluoromethyl)-benzyl)-1H-[1,2,3]triazol-4-yl]-pyridin-3-yl}-methanone (16), the selective, brain-penetrant NK1R antagonist used in this study. **(B)** Change from baseline for spontaneous alcohol cravings, as measured by twice-weekly ratings using the AUQ. The first value reflects the baseline rating obtained during the placebo lead-in week; the following ratings are from the active treatment phase. Data are means \pm SEM. Baseline values for placebo and LY686017 groups were 19.1 ± 2.4 and 19.7 ± 2.3 , respectively. Controlling for pretreatment baseline and sex, there was a highly significant decline of cravings over time ($F_{5,220} = 37.05$, $P < 0.0001$), and a significant main effect of treatment ($F_{1,100} = 4.4$, $P = 0.039$) on this outcome. **(C)** Change from baseline on weekly observer-based ratings using the Severity scale of the CGI



rating questionnaire. The first value reflects the baseline rating obtained during the placebo lead-in week; the following ratings are from the active phase. Average baseline ratings for the placebo and LY686017 groups were 3.80 ± 0.15 and 3.88 ± 0.13 , respectively. Data are means \pm SEM. Controlling for pretreatment baseline, sex, and body mass index, there was a significant effect of treatment ($F_{1,42} = 11.9$, $P = 0.001$). Very similar results were obtained on the improvement scale of the CGI (main treatment effect: $F_{1,44} = 8.4$, $P = 0.006$).

Finally, we investigated whether LY686017 affected brain responses to standardized affective stimuli, with a pattern suggestive of beneficial effects on relapse-related mechanisms. We measured blood oxygen–level–dependent (BOLD) activity using functional magnetic resonance imaging (fMRI), following presentation of negative and positive emotional stimuli from the International Affective Picture System (IAPS) (18) and pictures of alcoholic or neutral beverages (12). Alcoholics exhibit exaggerated behavioral and brain responses to images associated with negative affect, and conversely, exhibit reduced brain responses to standard positive images (19). In agreement with these earlier observations, we found that placebo-treated alcoholics showed robust brain responses to

negative affective images. Alcoholics who received LY686017 had less activation to the negative images than the placebo group in several brain regions associated with emotional response to visual stimuli [Fig. 4 and table S3 (12)]. In particular, the LY686017 group had less activation in the insula, a brain region whose activation correlates with subjective measures of craving (20), and which has recently been implicated in the maintenance of addictive behavior (21). In addition, the LY686017-treated group showed greater brain activation to the positive IAPS images than the placebo-treated group (Fig. 4). This may reflect an overall shift in the balance between positive and negative emotionality as indicated by the improvement detected on the CGI. A recent report suggests

that greater activation to positive images in the striatum and thalamus of treated alcoholics predicts less alcohol consumption over the next 6 months (22).

Basic neuroscience research has identified numerous candidate targets for pharmacological treatment of alcoholism (4), but translation into clinical development has been limited. Surrogate markers of efficacy that can be evaluated in alcoholics under safe, closely monitored conditions can facilitate translation. In this context, self-reported cravings are of interest, because they are triggered in humans by the same types of stimuli that, in animal models, induce relapse to alcohol-seeking (3, 5, 6). Human data also show that cravings correlate with clinical outcomes (23) and are sensitive to a clinically effective alcoholism medication, naltrexone (24). For these reasons, we chose spontaneous as well as challenge-induced cravings as primary outcomes in our study. Both these outcomes were beneficially affected by the NK1R antagonist.

The concept of craving and its ability to predict clinical outcomes has also invited some debate (25). Although effects on cravings and clinical outcomes do correlate in the case of naltrexone, it is unknown whether this relation will hold up for other pharmacodynamic mechanisms. This will only be possible to establish once additional compounds with clinical efficacy are identified. Until then, a translational approach that is guided by animal data and combines craving measures with a profile across a broader range of experimental outcomes appears to offer an attractive approach to drug development for alcoholism. Using this approach, we provide here consistent data across a range of measures suggesting that NK1R antagonism might be of therapeutic value in alcoholism. A possible mechanism for this is suppression of pathologically elevated amygdala activity thought to develop following a history of alcohol dependence (7). Our results were obtained in anxious alcoholics. Larger trials, stratified for anxiety measures, will be required to address whether the effect of NK1R antagonism is specific for this population.

Other NK1R antagonists have previously been tested clinically as a therapy for major depression (8, 26), but the results were inconsistent, and development of these drugs was stopped. Additional studies will be required to determine whether development of NK1R antagonists for alcoholism will be more successful, but recent findings with another stress-related neuropeptide system, CRH, suggest that this may be the case. Thus, analysis in animal models of alcoholism shows that the CRH system is quiescent under physiological or near-physiological conditions, and under these conditions, no activity of CRH antagonists is found. The CRH system is, however, pathologically activated following a history of alcohol dependence, revealing activity of CRH antagonists (7). These findings are in agreement with a general principle proposed for neuropep-

Fig. 3. Effect of the NK1R antagonist LY686017 on subjective craving and neuroendocrine response to a challenge session in hospitalized alcoholics.

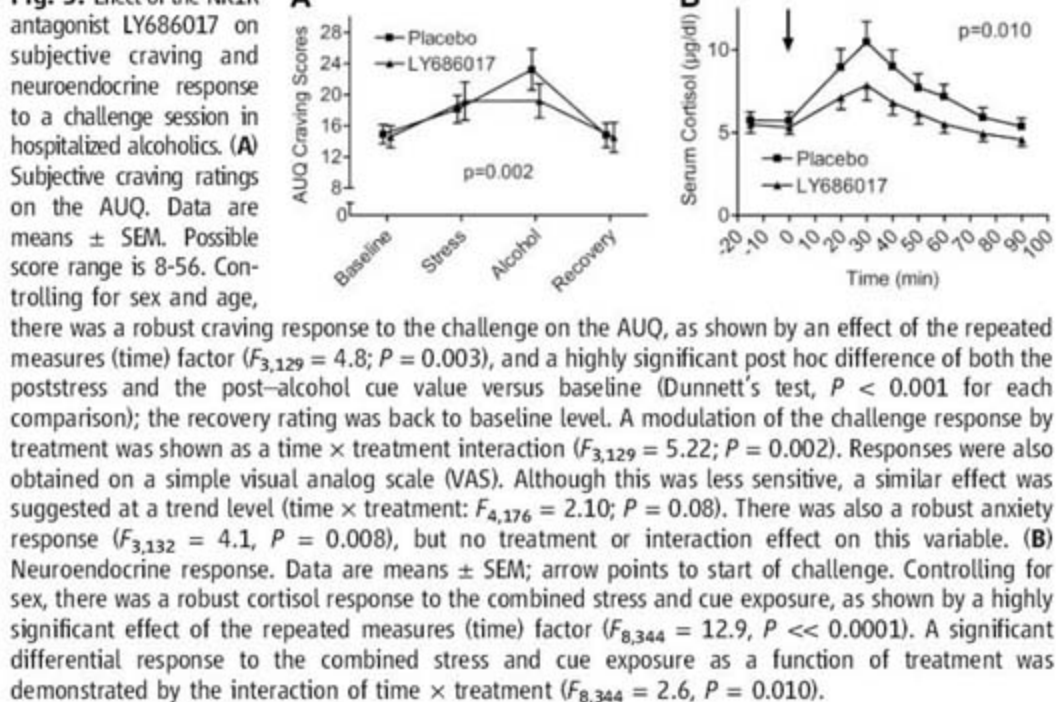
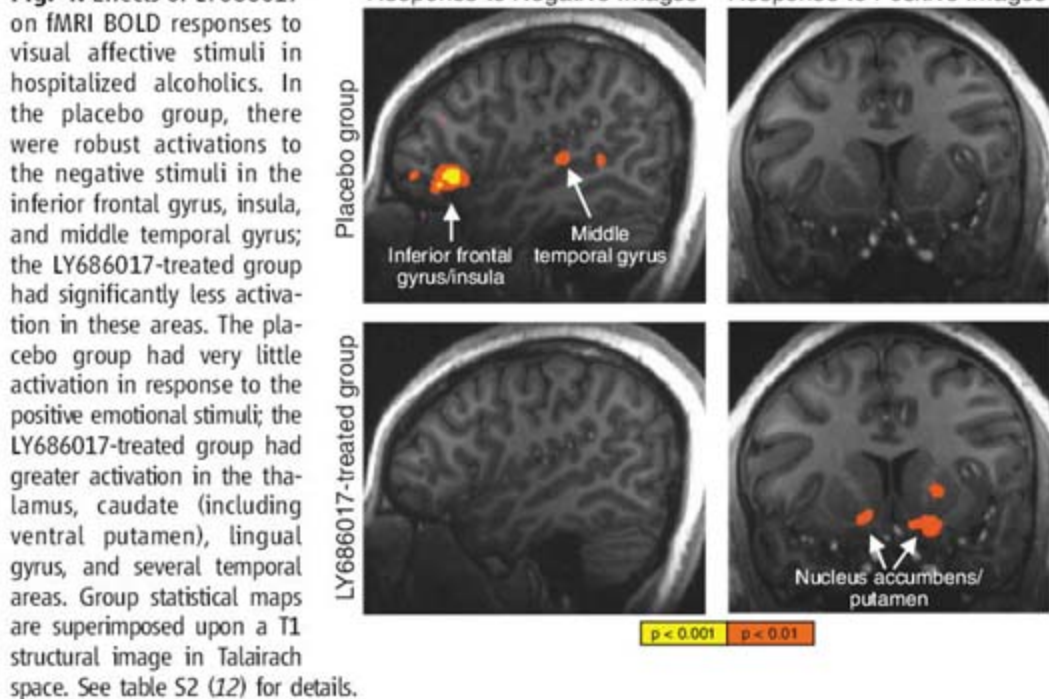


Fig. 4. Effects of LY686017 on fMRI BOLD responses to visual affective stimuli in hospitalized alcoholics.



tide systems (27). Along these lines, activation of the SP-NK1R system may not be a consistent feature of depressive illness. If, however, a pathological activation of the SP-NK1R system follows a history of alcohol dependence, similar to CRH, NK1 antagonism may have a considerable potential as a treatment for alcoholism.

References and Notes

- M. Ezzati, A. D. Lopez, A. Rodgers, S. van den Horn, C. J. Murray, *Lancet* **360**, 1347 (2002).
- K. D. Brownell, G. A. Marlatt, E. Lichtenstein, G. T. Wilson, *Am. Psychol.* **41**, 765 (1986).
- A. D. Lê et al., *Psychopharmacology (Berlin)* **135**, 169 (1998).
- M. Heilig, M. Egli, *Pharmacol. Ther.* **111**, 855 (2006).
- X. Liu, F. Weiss, *J. Neurosci.* **22**, 7856 (2002).
- R. Sinha, C. S. Li, *Drug Alcohol Rev.* **26**, 25 (2007).
- M. Heilig, G. F. Koob, *Trends Neurosci.* **30**, 399 (2007).
- A. Holmes, M. Heilig, N. M. J. Rupniak, T. Steckler, G. Griebel, *Trends Pharmacol. Sci.* **24**, 580 (2003).
- T. L. Ripley, C. A. Gadd, C. De Felipe, S. P. Hunt, D. N. Stephens, *Neuropharmacology* **43**, 1258 (2002).
- P. Murtra, A. M. Sheasby, S. P. Hunt, F. C. De, *Nature* **405**, 180 (2000).
- T. Furmark et al., *Biol. Psychiatry* **58**, 132 (2005).
- Materials and methods are available as supporting material on Science Online.
- C. De Felipe et al., *Nature* **392**, 394 (1998).
- J. C. Crabbe, T. J. Phillips, *Psychopharmacology (Berlin)* **174**, 539 (2004).
- M. A. Schuckit, T. L. Smith, J. Kalmijn, *Alcohol. Clin. Exp. Res.* **28**, 1449 (2004).
- A. K. Amegadzie et al., "Preparation of triazole derivatives as tachykinin receptor antagonists." (WO 2003091226 A1, World Intellectual Property Organization); published online 6 November 2003, www.wipo.int/pctdb/en/wo.jsp?wo=2003091226.
- C. D. Spielberger, R. L. Gorsuch, R. E. Lushene, *Manual for the State-Trait Anxiety Inventory* (Consulting Psychologist Press, Palo Alto, CA, 1970).
- P. J. Lang, M. M. Bradley, B. N. Cuthbert, *International Affective Picture System (IAPS): Technical Manual and Affective Ratings* (The Center for Research in Psychophysiology, Univ. of Florida, Gainesville, FL, 1995).
- J. M. Gilman, D. Hommer, *Addiction Biol.*, in press.
- A. L. Brody et al., *Arch. Gen. Psychiatry* **59**, 1162 (2002).
- N. H. Naqvi, D. Rudrauf, H. Damasio, A. Bechara, *Science* **315**, 531 (2007).
- A. Heinz et al., *Alcohol. Clin. Exp. Res.* **31**, 1138 (2007).
- B. A. Flannery, S. A. Poole, R. J. Gallop, J. R. Volpicelli, *J. Stud. Alcohol* **64**, 120 (2003).
- S. S. O'Malley, S. Krishnan-Sarin, C. Farren, R. Sinha, M. J. Kreek, *Psychopharmacology (Berlin)* **160**, 19 (2002).
- S. T. Tiffany, *Psychol. Rev.* **97**, 147 (1990).
- M. S. Kramer et al., *Science* **281**, 1640 (1998).
- J. M. Lundberg, T. Hökfelt, *Trends Neurosci.* **6**, 325 (1983).
- We thank nurse manager J. Johnson and her treatment staff at the 1SE Unit of the NIH Clinical Center, as well as D. Hill, L. Doty, and C. Jones for their invaluable contributions to this study. R. Anton, S. Thomas, R. Miranda, and P. Monti are gratefully acknowledged for input in the design stages of the clinical study. Supported by the intramural research budget of the National Institute on Alcohol Abuse and Alcoholism, NIH. This work was carried out under a Collaborative Research and Development Agreement (CRADA) between the U.S. Government and Eli Lilly and Co. All coauthors so indicated are employees of Eli Lilly and Co.; the remaining authors have no conflict of interest to disclose.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1153813/DC1

Materials and Methods

Fig. S1

Tables S1 to S3

References

5 December 2007; accepted 6 February 2008

Published online 14 February 2008;

10.1126/science.1153813

Include this information when citing this paper.

Using Engineered Scaffold Interactions to Reshape MAP Kinase Pathway Signaling Dynamics

Caleb J. Bashor,^{1,2} Noah C. Helman,¹ Shude Yan,¹ Wendell A. Lim^{1*}

Scaffold proteins link signaling molecules into linear pathways by physically assembling them into complexes. Scaffolds may also have a higher-order role as signal-processing hubs, serving as the target of feedback loops that optimize signaling amplitude and timing. We demonstrate that the Ste5 scaffold protein can be used as a platform to systematically reshape output of the yeast mating MAP kinase pathway. We constructed synthetic positive- and negative-feedback loops by dynamically regulating recruitment of pathway modulators to an artificial binding site on Ste5. These engineered circuits yielded diverse behaviors: ultrasensitive dose response, accelerated or delayed response times, and tunable adaptation. Protein scaffolds provide a flexible platform for reprogramming cellular responses and could be exploited to engineer cells with novel therapeutic and biotechnological functions.

In cells, signaling proteins that make up a pathway are often physically organized into complexes by scaffold proteins (1–3). Scaffolds direct information flow; they promote signaling between proper protein partners and prevent improper cross talk. Scaffolds may also play a role in shaping the quantitative response behavior of a pathway. The scaffold complex could serve as a central hub for feedback loops that modulate the recruitment or activity of

pathway members on the scaffold. Such feedback loops could tune pathway dose response and dynamics—the change in output over time. Quantitative response behavior is critical for signaling; the behavior of a pathway must match its specific physiological function (4). Scaffolds may therefore provide a platform for evolutionarily tuning response behaviors for optimal fitness (5, 6).

We used a synthetic biology approach to explore this hypothesis; we tested whether a scaffold protein can be used as a platform for engineering synthetic feedback loops and whether these loops can be used to systematically reshape pathway response behavior (7). We used the yeast mating mitogen-activated protein (MAP) kinase pathway as a model system, because it is highly tractable for pathway engineer-

ing. First, proper connectivity of this pathway is dependent on the scaffold protein Ste5, which binds the three core kinases—Ste11 (a MAP kinase kinase kinase or MAPKKK), Ste7 (a MAP kinase kinase, or MAPKK), and Fus3 (a MAP kinase, or MAPK)—that successively phosphorylate and activate one another (Fig. 1A) (8, 9). The critical role in determining pathway connectivity is highlighted by the observation that chimeric scaffolds can be used to redirect pathway input and output linkages (10, 11). Second, MAP kinase pathways appear to be functionally plastic; they are found in all eukaryotic species, but in individual cases display widely varied behaviors. For example, the yeast mating pathway shows a largely linear transcriptional response (12–14), whereas the *Xenopus* oocyte maturation pathway displays a switchlike dose response (15). MAPK pathways also show diverse dynamic behavior; some yield a sustained response to stimulation, whereas others show a pulselike transient response. These distinct pathway dynamics are critical for determining physiological output (16–21).

Our goal was to overlay the endogenous mating pathway with synthetic feedback loops in order to systematically alter its response to mating pheromone (α -factor) stimulation. A simple way to construct a synthetic feedback loop would be to dynamically recruit pathway modulators to the scaffold in a manner that is dependent on pathway output. We first tested whether constitutive recruitment of modulator proteins could alter pathway flux. We created a new recruitment site on Ste5 by fusing a leucine zipper heterodimerization module (22) to its C terminus. Modulator proteins fused to complementary zippers were expressed and recruited to the scaffold (Fig. 1B). Two pathway modulators were recruited:

¹Department of Cellular and Molecular Pharmacology, University of California at San Francisco, 600 16th Street, San Francisco, CA 94158, USA. ²Graduate Group in Biophysics, University of California at San Francisco, 600 16th Street, San Francisco, CA 94158, USA.

*To whom correspondence should be addressed. E-mail: lim@cmp.ucsf.edu

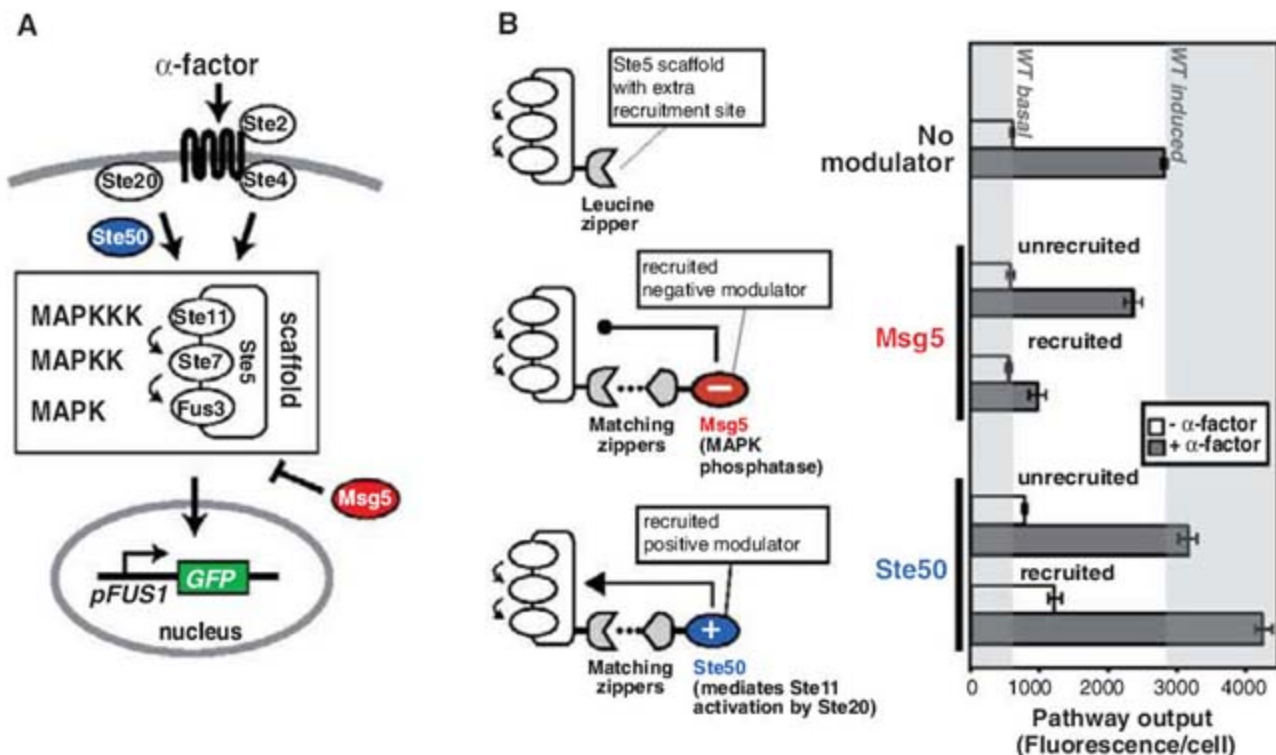
Ste50 and Msg5 (Fig. 1A). Ste50 is a positive modulator, an adaptor that promotes interaction of the MAPKKK Ste11 with its upstream activator, the p21-activated protein kinase (PAK)-like kinase, Ste20 (23, 24). Msg5 is a negative modulator, a MAPK phosphatase that inactivates phosphorylated Fus3 MAPK (25, 26). When artificially recruited to the Ste5 scaffold through a

leucine zipper interaction, Msg5 and Ste50 showed strong but opposite effects on pathway output, measured using a fluorescent transcriptional reporter [see supporting online text and (27)]. Recruitment of the positive modulator (Ste50) increased the steady-state output of the activated pathway, whereas recruitment of the negative modulator (Msg5) decreased output,

nearly eliminating any input-stimulated response. Unrecruited Ste50 and Msg5 had much smaller effects when expressed at the same level. Thus, the impact of modulators on pathway flux was enhanced by recruitment to the scaffold.

To build synthetic feedback loops, we then placed the modulators under the control of a mating-dependent promoter (*pFIG1*) (Fig. 2A),

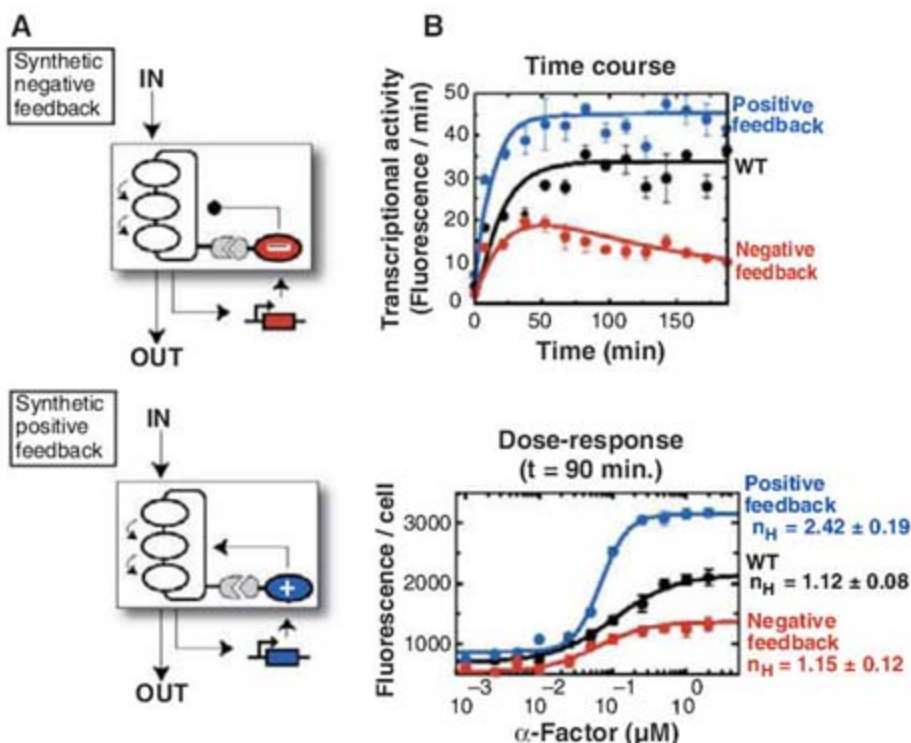
Fig. 1. Tuning output from yeast mating MAPK pathway by artificially recruiting positive or negative modulators to Ste5 scaffold protein. (A) Yeast mating pathway: α -factor activates receptor, Ste2, and G_{β} subunit, Ste4; activated Ste4 recruits Ste5 complex to membrane, allowing PAK-like kinase Ste20 (membrane-localized) to activate MAPKKK Ste11; Ste11 and downstream kinases, Ste7 (MAPKK) and Fus3 (MAPK), are colocalized on the scaffold; activation of cascade leads to transcriptional program [reporter: Fus1 promoter-GFP (green fluorescent protein)]. We have used pathway modulators outside of core cascade: Ste50 (positive, blue) promotes activation of Ste11 by Ste20; Msg5 (negative, red) is MAPK phosphatase that deactivates Fus3. (B) Synthetic recruitment of pathway modulators to Ste5 scaffold via leucine zipper interaction. A basic zipper was fused to Ste5, and the complementary acidic zipper was fused to modulators (zipper $K_d = 6.1$ nM; see supporting online text). Control ("unrecruited") modulators were fused to noncomplementary zipper. Pathway output was assessed via *pFUS1*-GFP expression. Basal output is before



stimulation; induced output is 120 min after stimulation with 2 μ M α -factor (saturating). Error bars represent standard deviation of three experiments. Strains were *ste5* Δ with integrated Ste5-zipper fusion expressed from *pSTE5* promoter. Modulator-zipper fusions were integrated and expressed from *pCYC1* promoter. Reversed orientation of zippers (fig. S2E) gave similar results.

stimulation; induced output is 120 min after stimulation with 2 μ M α -factor (saturating). Error bars represent standard deviation of three experiments. Strains were *ste5* Δ with integrated Ste5-zipper fusion expressed from *pSTE5* promoter. Modulator-zipper fusions were integrated and expressed from *pCYC1* promoter. Reversed orientation of zippers (fig. S2E) gave similar results.

Fig. 2. Building synthetic feedback loops by dynamically regulating recruitment of modulators to the Ste5 scaffold. (A) Negative- and positive-feedback loop design. Modulator-zipper fusions (negative, Msg5; positive, Ste50) are expressed from a mating-responsive promoter (*pFIG1*). Stimulation induces expression of the modulator, which is then recruited to the Ste5-complex (boxed), where it can modulate pathway flux. (B) Behavior of synthetic feedback loops. All circuits were in *far1* Δ strain (designated "WT"), which does not undergo mating-induced cell cycle arrest. Thus, cells are uniform in size, facilitating fluorescence-activated cell sorting analysis. Negative-feedback circuit (red) shows an initial rate of activation similar to the wild-type, but peaks at ~ 35 min, and adapts to steady-state output about one-third that of wild-type. Positive-feedback circuit (blue) shows dynamics similar to wild type's, but with steady-state output ~ 1.5 times as great as wild type's. The dose-response curves (bottom) show that the positive-feedback circuit displays more switchlike activation (apparent Hill coefficient $n_H = 2.42 \pm 0.19$ versus wild-type, $n_H = 1.12 \pm 0.08$). Points represent mean values for three experiments \pm standard deviation. Pathway is activated with 2 μ M α -factor. Best-fit lines generated as described in supporting online text.



so that they were only expressed upon pathway activation. The positive-feedback circuit (using Ste50) increased steady-state pathway output and led to a more switchlike dose response (apparent Hill coefficient n_H increased from 1.12 to 2.42). Similar behavior was generated using positive-feedback loops in which constitutive pathway alleles were expressed from a mating-responsive promoter (28).

The engineered negative-feedback circuit (using Msg5) displays adaptation. The cells initially responded like wild-type, but after 35 min, showed a decrease in output, even with continued stimulation. Adaptation is critical for homeostatic and sensing systems. It can be important to limit the magnitude and duration of an output that is harmful or has a high metabolic cost. Adaptation is also used in sensing systems (e.g., vision or chemotaxis) that automatically desensitize to a continuous stimulus, allowing for detection of input changes over a large dynamic range (29, 30).

One advantage of engineered feedback loops is the ability to systematically explore how the alteration of specific circuit parameters affects pathway behavior. Simulations indicate that adaptation in the simple negative-feedback circuit can be tuned by adjusting feedback strength (Fig. 3A). We explored two methods for adjusting feedback strength. First, we changed the strength of the leucine zipper interaction used to recruit Msg5 (Fig. 3B), using a set of three leucine zipper pairs that bind with the following affinities: $K_d = 6, 40,$ and 810 nM (22). Second, we changed the strength of the promoter con-

trolling expression of recruited Msg5 (Fig. 3C), using a pair of mating promoters: *pFIG1* (strong) and *pPRM2* (weak) (see fig. S4). As predicted, either method results in a decrease in the steady-state pathway output.

Because these synthetic feedback circuits rely on modulator recruitment, they could be regulated by competitive interactions that block recruitment. We tested whether competitors could be used to build more complex negative-feedback circuits that displayed a pulslike activation response (high maximal output followed by low steady-state output) (Fig. 4A). We constitutively expressed a decoy leucine zipper that competes with the scaffold protein (Ste5-zipper) for binding to the negative modulator (Msg5-zipper). Because the decoy has a higher affinity, it initially acts as a sink; after pathway activation, newly expressed negative modulator bound to the decoy zipper, preventing recruitment to the scaffold. Only after the decoy zipper is saturated is additionally expressed negative modulator recruited to the scaffold, which results in pathway repression. Indeed, this delayed negative-feedback loop led to a pulslike response (higher maximal output, followed by decrease in output). Moreover, the sharpness of the response could be modulated by adjusting the level at which the decoy zipper was expressed; higher decoy expression led to a more pulslike response.

The interplay of competing scaffold interactions and variable expression can be used to generate additional dynamic behaviors, including systems with faster or slower response time. The

rise time of a pathway—how fast a response occurs after input—can be critical for function. A pathway that detects a toxic stress may require a fast response. A delayed response may be required if the response is energetically very costly or if there is a high level of input noise (a delay circuit could prevent misactivation by transient input fluctuations, while still allowing activation by a sustained input).

We were able to alter the mating pathway to show accelerated response time, while maintaining a wild-type level of maximal pathway output (Fig. 4B). In this accelerator circuit, the positive modulator (Ste50-zipper) was constitutively expressed, but the negative modulator (Msg5-zipper, high affinity) was inducibly expressed. This result supports the idea that negative feedback can speed the time it takes to reach steady-state (31), albeit at a reduced steady-state output. But here, the wild-type magnitude of output is achieved by the added presence of the positive modulator.

We also generated a delay circuit by constitutively expressing a negative modulator (Msg5-zipper) and inducibly expressing a high-affinity decoy zipper (Fig. 4C). The pathway initially showed a weak response to stimulation because the recruited negative modulator kept the pathway off. After a delay of ~ 50 min, however, a small but sufficient level of pathway activation was reached, and expression of the high-affinity decoy zipper displaced the negative modulator, which allowed pathway activation.

Competition between positive and negative modulators can also be used to alter dose-response

Fig. 3. Strength of synthetic negative-feedback circuit can be tuned by either altering recruitment affinity or inducible expression level of negative modulator. (A) Computational model (see supporting online text) predicts that adjustments in negative-feedback gain should tune adaptation behavior. (B) Adjusting feedback gain by varying modulator recruitment strength. Three affinity variants of the leucine zipper (strong, $K_d = 6.1$ nM; medium, $K_d = 41$ nM; and weak, $K_d = 810$ nM) were fused to Msg5 (negative modulator). Promoter controlling modulator expression was constant (*pPRM2*). (C) Adjusting feedback gain by varying modulator promoter strength. Msg5-zipper is expressed from two mating-induced promoters: strong, *pFIG1*, and weak, *pPRM2* (fig. S4). Zipper fused to Msg5 was constant (weak). Orange plots in (B) and (C) show same strain (circuit with weak zipper and weak promoter). Points represent mean values for three experiments \pm standard deviation.

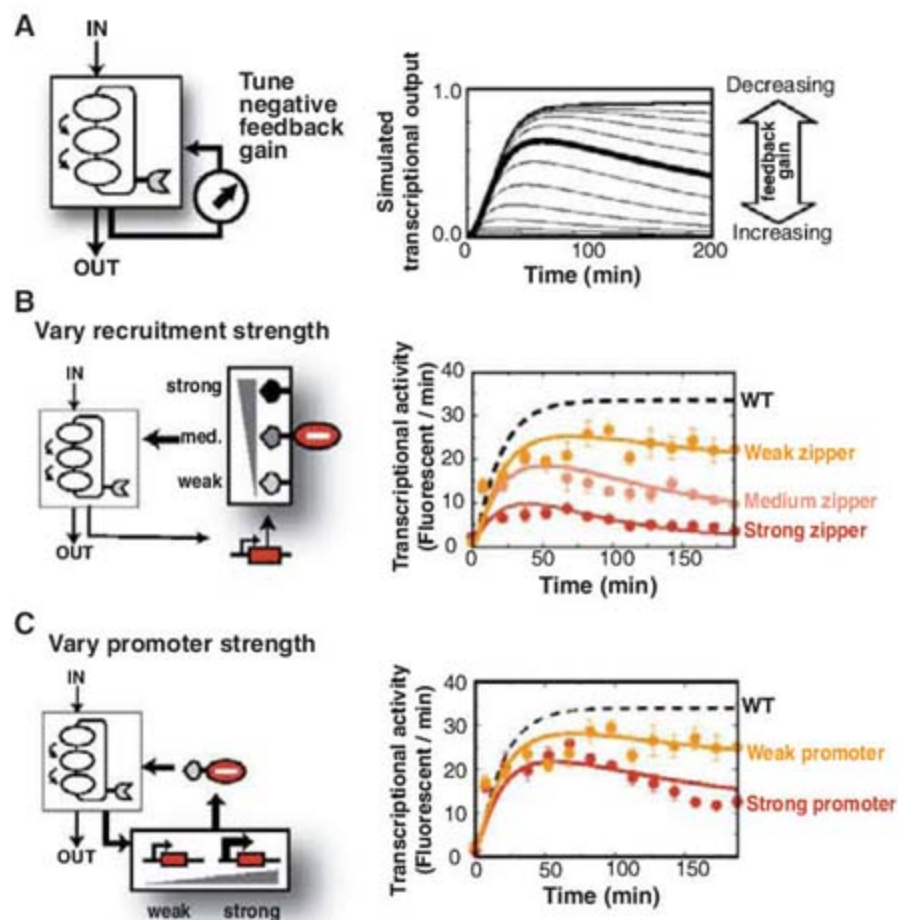
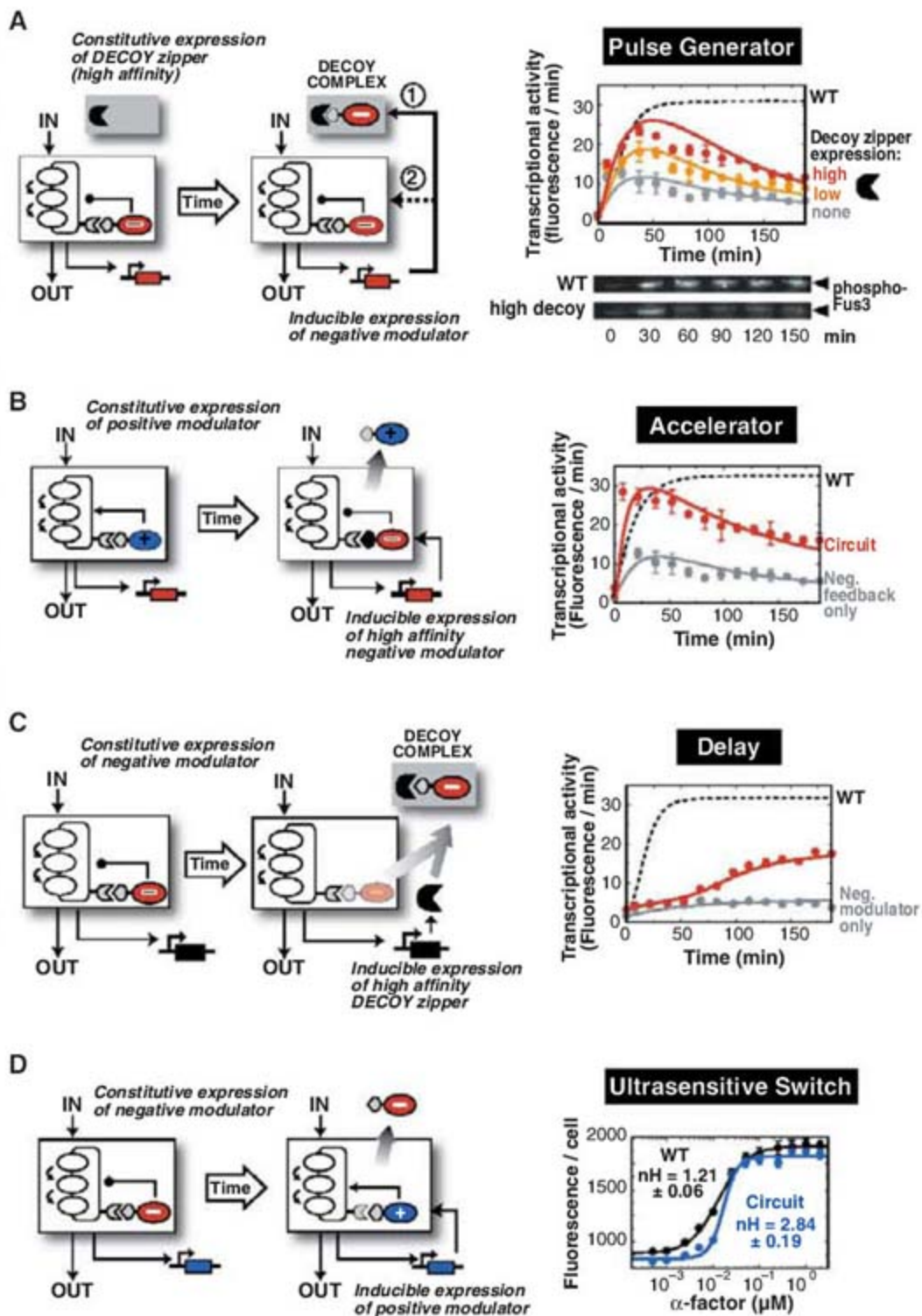


Fig. 4. Recruitment-based tool kit can be used to engineer diverse response behaviors. **(A)** Pulse behavior is built from simple negative-feedback loop (Fig. 2) by adding constitutively expressed decoy zipper (GST-zipper) that competes with the Ste5-zipper. Negative modulator (Msg5-zipper) is complementary to both Ste5 and decoy zippers, but decoy zipper binds with higher affinity ($K_d = 6.1$ versus 41 nM). When pathway is induced and negative modulator is expressed, decoy initially acts as binding sink. Only after decoy is saturated will modulator bind to scaffold and repress pathway flux. Circuit behavior monitored by *pFUS1*-GFP reporter (top) and antibody against phospho-Fus3 Western blot (bottom). **(B)** Accelerator circuit is generated by constitutively expressing positive modulator (promoter, *pSTE5*; gene, Ste50-zipper) combined with the simple negative-feedback loop (Fig. 2). Circuit reaches the maximal output observed for the wild-type circuit in <20 min (versus WT: ~75 min). **(C)** Delay response is generated by constitutively expressing negative modulator (promoter, *pSTE5*; gene, Msg5-zipper) and inducibly expressing decoy zipper that binds the negative effector-zipper (promoter, *pFIG1*; gene, GST-zipper) with higher affinity than the Ste5-zipper ($K_d = 6.1$ versus 41 nM). Negative modulator represses pathway until sufficient decoy zipper is expressed to relieve repression. **(D)** Enhanced ultrasensitive switch is built by constitutively expressing negative modulator (promoter, *pSTE5*; gene, Msg5-zipper) combined with the simple positive-feedback loop (Fig. 2). Negative modulator is displaced by inducibly expressed positive modulator (both have same zipper; $K_d = 41$ nM). Dose-response analysis shows increased ultrasensitivity (apparent Hill coefficient $n_H = 2.84 \pm 0.19$ versus WT, $n_H = 1.21 \pm 0.06$; simple positive-feedback loop, $n_H = 2.42$). See supporting online text for details on construction and analysis of these and related circuits.



behavior. We built a circuit with enhanced ultrasensitive switch behavior by constitutively expressing a negative modulator (Msg5-zipper) and inducibly expressing a positive modulator (Ste50-zipper) (Fig. 4D). This circuit is a double-positive-feedback loop—induced expression of Ste50-zipper directly increased pathway output, but also relieved the inhibitory effect of Msg5-zipper by displacing it from the scaffold. The dose-response profile for the resulting circuit showed an increase in cooperativity.

We have used a simple principle, recruitment of pathway modulators to a scaffold (32), to systematically alter a single MAPK pathway so that it displays a wide range of quantitative response behaviors. The evolutionary diversifi-

cation afforded by scaffolds may explain their common use in signaling pathways (33–37). The success of this simple recruitment-based engineering strategy suggests that it may be possible to reprogram cellular responses with high precision. In mammalian and plant cells, MAPK cascades play a central role in regulating differentiation, proliferation, apoptosis, immune response, and stress responses. Thus, the ability to tune MAPK signaling may facilitate the engineering of cells with novel therapeutic or biotechnological functions.

References and Notes

- W. R. Burack, A. S. Shaw, *Curr. Opin. Cell Biol.* **12**, 211 (2000).
- T. Pawson, J. D. Scott, *Science* **278**, 2075 (1997).

- R. P. Bhattacharyya, A. Remenyi, B. J. Yeh, W. A. Lim, *Annu. Rev. Biochem.* **75**, 655 (2006).
- B. N. Kholodenko, *Nat. Rev. Mol. Cell Biol.* **7**, 165 (2006).
- J. W. Locasale, A. S. Shaw, A. K. Chakraborty, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 13307 (2007).
- J. E. Ferrell, *Sci. STKE* **2000**, PE1 (2000).
- Materials and methods are available as supporting material on Science Online.
- K. Y. Choi, B. Satterberg, D. M. Lyons, E. A. Elion, *Cell* **3**, 499 (1994).
- J. A. Printen, G. F. Sprague, *Genetics* **138**, 609 (1994).
- S. H. Park, A. Zarrinpar, W. A. Lim, *Science* **299**, 1061 (2003).
- K. Harris *et al.*, *Curr. Biol.* **11**, 1815 (2001).
- M. A. Poritz, S. Malmstrom, M. K. Kim, P. J. Rossmeissl, A. Kamb, *Yeast* **18**, 1331 (2001).
- A. Colman-Lerner *et al.*, *Nature* **437**, 699 (2005).
- S. Paliwal *et al.*, *Nature* **446**, 46 (2007).
- J. E. Ferrell Jr., E. M. Machleder, *Science* **280**, 895 (1998).

16. C. J. Marshall, *Cell* **80**, 179 (1995).
 17. S. Sasagawa, Y. Ozaki, K. Fujita, S. Kuroda, *Nat. Cell Biol.* **7**, 365 (2005).
 18. L. O. Murphy, J. Blenis, *Trends Biochem. Sci.* **31**, 268 (2006).
 19. M. Villedieu *et al.*, *Gynecol. Oncol.* **101**, 507 (2006).
 20. B. K. Choi, C. H. Choi, H. L. Oh, Y. K. Kim, *Neurotoxicity* **25**, 915 (2004).
 21. S. D. Santos, P. J. Verwee, P. I. Bastiaens, *Nat. Cell Biol.* **9**, 247 (2007).
 22. A. Acharya, S. B. Ruvinov, J. Gal, C. Vinson, *Biochemistry* **41**, 14122 (2002).
 23. M. Ramezani-Rad, *Curr. Genet.* **43**, 161 (2003).
 24. C. Wu, E. Leberer, D. Y. Thomas, M. Whiteway, *Mol. Biol. Cell* **10**, 2425 (1999).
 25. X. L. Zhan, R. J. Deschenes, K. L. Guan, *Genes Dev.* **11**, 1690 (1997).
 26. J. Andersson, D. M. Simpson, M. Qi, Y. Wang, E. A. Elion, *EMBO J.* **23**, 2564 (2004).
 27. R. P. Bhattacharyya *et al.*, *Science* **311**, 822 (2006).
 28. N. T. Ingolia, A. W. Murray, *Curr. Biol.* **17**, 668 (2007).
 29. N. Barkai, S. Leibler, *Nature* **387**, 913 (1997).
 30. B. Alberts *et al.*, *Essential Cell Biology* (Garland Science, London, ed. 2, 2003).
 31. N. Rosenfeld, M. B. Elowitz, U. Alon, *J. Mol. Biol.* **323**, 785 (2002).
 32. M. Ptashne, A. Gann, *Genes and Signals* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001).
 33. D. C. Popescu, A. J. Ham, B. H. Shieh, *J. Neurosci.* **26**, 8570 (2006).
 34. K. Scott, C. S. Zuker, *Nature* **395**, 805 (1998).
 35. F. D. Smith, L. K. Langeberg, J. D. Scott, *Trends Biochem. Sci.* **31**, 316 (2006).
 36. S. C. Strickfaden *et al.*, *Cell* **128**, 519 (2007).
 37. P. Mishra *et al.*, *Cell* **131**, 80 (2007).
 38. We thank H. El-Samad, T. Kortemme, H. Madhani, C. Tang, C. Voigt, J. Weissman, and the Lim laboratory for advice and comments. Supported by University of

California GREAT fellowship (C.J.B.), American Cancer Society Postdoctoral Fellowship (N.C.H.), Jane Coffin Childs Fellowship (S.Y.) and grants from the U.S. Defense Advanced Research Projects Agency (Biological Input/Output Systems); NIH Nanomedicine Development Centers (Roadmap); National Institute of General Medical Science, NIH; Packard Foundation; and Rogers Family Foundation (W.A.L.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/319/5869/1539/DC1
 Materials and Methods

SOM Text

Figs. S1 to S6

Tables S1 to S3

References

1 October 2007; accepted 11 February 2008
 10.1126/science.1151153

Synaptic Theory of Working Memory

Gianluigi Mongillo,^{1*†} Omri Barak,^{2*} Misha Tsodyks^{2‡§}

It is usually assumed that enhanced spiking activity in the form of persistent reverberation for several seconds is the neural correlate of working memory. Here, we propose that working memory is sustained by calcium-mediated synaptic facilitation in the recurrent connections of neocortical networks. In this account, the presynaptic residual calcium is used as a buffer that is loaded, refreshed, and read out by spiking activity. Because of the long time constants of calcium kinetics, the refresh rate can be low, resulting in a mechanism that is metabolically efficient and robust. The duration and stability of working memory can be regulated by modulating the spontaneous activity in the network.

Working memory (WM) enables the temporary holding of information for processing purposes, playing a crucial role in the execution of a wide range of cognitive tasks (1). In the delayed-response paradigm, a stimulus that is briefly presented to an animal has to be kept for several seconds until the execution of a task. Enhanced, stimulus-specific spiking activity has been observed during this delay period and is considered to be a neuronal correlate of WM (2–5). The current theoretical framework holds that delay activity emerges either from intrinsic cell properties (6, 7) or as persistent reverberations in selective neural populations coding for different memories (8–12). These populations are formed during learning via long-term synaptic modifications (13). However, electrophysiological studies have shown that the delay activity increase can be very modest (14, 15), sometimes disappearing completely during part

of the delay period (16). These observations suggest that WM might not reside entirely in the spiking activity. Furthermore, holding information in a spiking form is energetically expensive because of the high metabolic cost of action potentials (17). Here, we present an alternative account based on properties of excitatory synaptic transmission in the prefrontal cortex (PFC) (18). The PFC is a cortical area implicated in WM (4), and excitatory synaptic transmission in this area can be markedly facilitatory, unlike sensory areas where it is mostly depressing (19, 20). We therefore propose that an item is maintained in the WM state by short-term synaptic facilitation mediated by increased residual calcium levels at the presynaptic terminals of the neurons that code for this item (21). Because removal of residual calcium from presynaptic terminals is a relatively slow process (22, 23), the memory can be transiently held for about 1 s without enhanced spiking activity.

We implemented this mechanism with a recurrent network of integrate-and-fire neurons (24). The network encodes a set of memories (items) by randomly composed selective populations of excitatory neurons (Fig. 1B). Connections between the neurons coding for the same memory are stronger than connections between different populations, mimicking the result of prior long-term Hebbian learning (25) or intrinsic clustering of recurrent connections (26). Inhibitory neurons are connected to the excitatory

ones in a nonstructured way, resulting in competition between different memories [see supporting online material (SOM)]. All excitatory-to-excitatory connections display facilitating transmission, described by a phenomenological model of short-term plasticity (20, 27). Synaptic efficacy is modulated by the amount of available resources (x , normalized so that $0 < x < 1$) and the utilization parameter (u) that defines the fraction of resources used by each spike, reflecting the residual calcium level (22, 23) (Fig. 1A and SOM). Upon a spike, an amount ux of the available resources is used to produce the postsynaptic current, thus reducing x . This process mimics neurotransmitter depletion. The spike also increases u , mimicking calcium influx into the presynaptic terminal and its effects on release probability. Between spikes, x and u recover to their baseline levels ($x = 1$ and $u = U$) with time constants τ_D (depressing) and τ_F (facilitating), respectively. The phenomenological model reproduces the behavior of cortical synapses, both depressing ($\tau_D > \tau_F$) and facilitating ($\tau_F > \tau_D$) (27). For PFC facilitating excitatory connections, the experimental fit reports $\tau_F \gg \tau_D$ (18), with τ_F on the order of 1 s.

The simulations begin with loading one item into WM by providing transient external excitation to the corresponding neural population (Fig. 2A). The population activity increases for the duration of the input, changing the internal state of the synaptic connections. The connections are both depressed (reduced x) and facilitated (increased u), with depression dominant on the time scale of τ_D and facilitation dominant on the time scale of τ_F (where $\tau_D = 0.2$ s and $\tau_F = 1.5$ s; see SOM for all parameter values). As long as the synapses remain facilitated, the memory can be reactivated by presenting a weak nonspecific excitatory input to the whole network (gray shading), even though the neural activity is at the spontaneous level. Reactivation is expressed as a short epoch of synchronized activity [“population spike” (PS)], where almost every neuron in the population fires a spike within an interval of about 20 ms (28, 29). Even though the reactivating signal is nonspecific (that is, it uniformly

¹Group for Neural Theory, Département d’Etudes Cognitives, Ecole Normale Supérieure et Collège-de-France, Paris, France.

²Department of Neurobiology, Weizmann Institute of Science, Rehovot, Israel.

*These authors contributed equally to this work.

†Present address: Laboratoire de Neurophysique et Physiologie, Université Paris-Descartes, CNRS-UMR8119, and Franco-Israeli Laboratory of System Neurophysiology and Neurophysics, Paris, France.

‡CNRS visiting member of the Group for Neural Theory, Ecole Normale Supérieure et Collège-de-France.

§To whom correspondence should be addressed. E-mail: misha@weizmann.ac.il

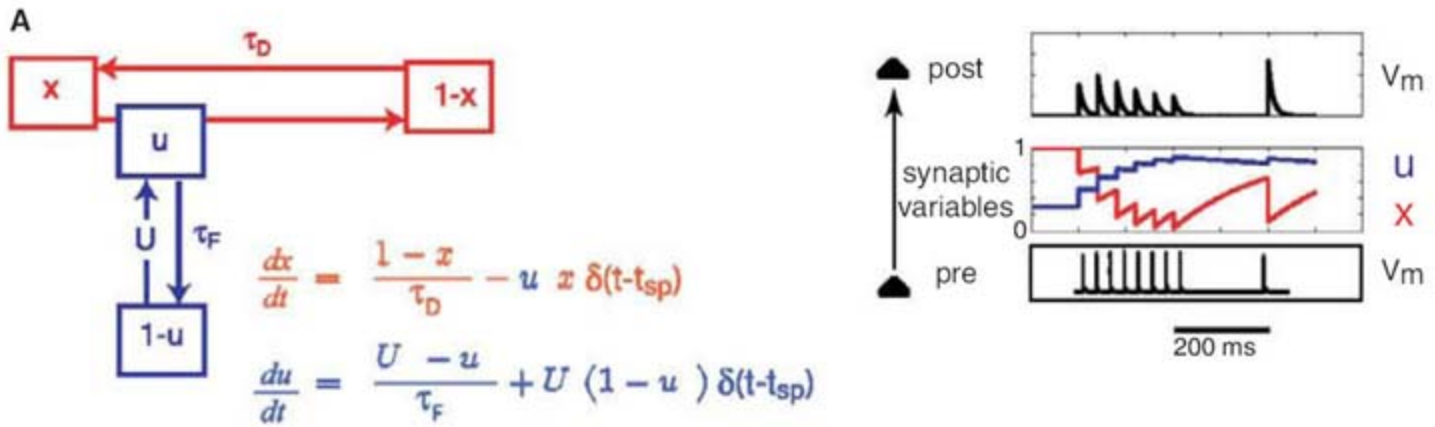
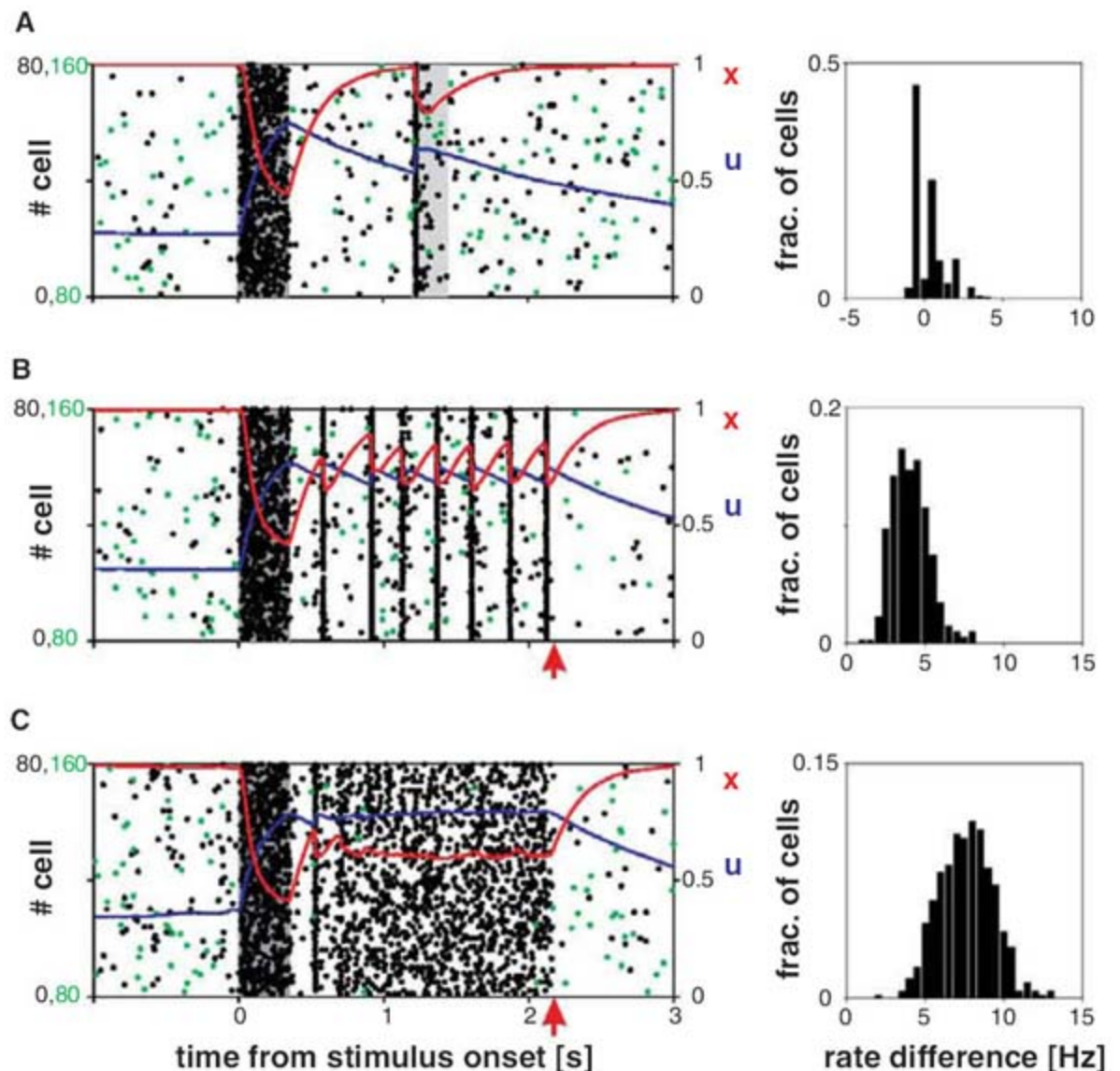


Fig. 1. Physiology and anatomy of the network. **(A)** Short-term synaptic plasticity model. (Left) Kinetic scheme with the corresponding equations for synaptic variables. δ , Dirac delta function; t_{sp} , time of presynaptic spike. (Right) Example of the postsynaptic response to a train of presynaptic action potentials in the case of a facilitating connection. During the train, u increases (facilitation) and x decreases (depression). Synaptic efficacy is modulated by the product ux . v_m , membrane potential. **(B)** Network architecture. Colored triangles are excitatory neurons that code for different memories. Black open triangles are nonselective excitatory neurons. Black circles are inhibitory neurons with nonstructured connections to the entire network.

Fig. 2. Memory maintenance with synaptic facilitation. **(A)** The item is loaded into the memory by activating the corresponding population at time $t = 0$ (dark shading). A nonspecific read-out signal is applied to the entire excitatory network as explained in the text (gray shading), leading to the selective reactivation of the target population via the PS. Black and green dots, spike rasters from a subset (10%) of neurons from the targeted and one of the nontargeted populations, respectively; red curve, average value of x in the synaptic connections of the target population; blue curve, average value of u for the same synapses. **(B)** Same as **(A)**, with an increased background input. The target population reactivates spontaneously with a set of PSs. Red arrow indicates termination of excitation. **(C)** A further increase in background input leads to WM with asynchronous elevated firing in the target population. (Right) Corresponding histograms for **(A)** to **(C)** of the difference in firing rate between the delay period and the spontaneous state for different neurons in the target population [note the different x axis in **(A)**]. The delay period is defined in **(A)** as the interval after the termination of the selective input until the onset of the read-out signal and in **(B)** and **(C)** as the period until the decrease of external excitation.



targets all the neurons), the network response is memory-specific: The neurons coding for the loaded item produce a PS; the others stay at baseline activity level. The PS also refreshes the memory by producing additional facilitation, thus enabling subsequent memory reactivations. In the absence of reactivating signals, the memory fades away over a time scale on the order of τ_F .

In the above scenario, the network has a single stable activity state corresponding to the spontaneous activity, thus appropriately timed external signals are required to extract the memory from synaptic to spiking form. A more persistent form of WM requires the selective population to exhibit a bistable activity regime, where the spontaneous state coexists with another stable state (8). Our network can be forced into this regime by increasing spontaneous activity by means of a global nonspecific background input (see SOM for the mathematical analysis). Accordingly, we simulated the network for increasing levels of background input. In the bistable regime, PSs become persistent without reactivating inputs (Fig. 2B). Each reactivation increases u and

decreases x , the latter terminating the PS. The time between subsequent PSs is controlled by the recovery from synaptic depression so that the PSs tend to occur with a period on the order of τ_D . With a τ_D compatible with (18), this would correspond to cortical oscillations in the theta-range, as observed during WM experiments (30, 31). Because $\tau_F \gg \tau_D$, the decay of the utilization factor is balanced by the increase produced by the PSs, so that u remains at sufficiently high levels for subsequent PSs to emerge. Persistent PSs can be terminated by reducing background input, thus restoring the network to the transient regime. A different model of persistent PSs is based on the increase in asynchronous transmitter release (32). In the simulation presented in Fig. 2B, neurons coding for a given memory exhibit highly coherent firing during the PSs. In more realistic networks, PSs could be broader and consist of random subpopulations of neurons, resulting in less pronounced synchrony (fig. S2). If nonspecific background input increases further, the network exhibits bistability between a spontaneous state

and an asynchronous state of enhanced firing rate (Fig. 2C). Information about the memory is maintained in both synaptic and spiking form. This regime exists for sufficiently strong recurrent connections, which could possibly result from extensive learning.

The use of residual calcium at synaptic terminals as a memory "buffer" requires low emission rates (Fig. 2, histograms). Moreover, the buffer content is not substantially affected by the neural activity in the rest of the cortex. When we presented the network with a noisy input that targeted a random subset of excitatory neurons for a brief duration, the increased firing of neurons receiving the input suppressed the memory-related spiking activity (Fig. 3, teal shading). The information, however, remained in the increased utilization factor in the memory population; hence, the spiking activity resumed after the termination of the suppressing input. The same feature enables the network to keep multiple items simultaneously with interleaved PSs (33). We illustrated this possibility in two different conditions: (i) when the network has a single activity state and PSs result from a sequence of reactivating signals (Fig. 3A) or (ii) when the network exhibits persistent PSs (Fig. 3B). When a new item is presented, the previous memory is temporarily suppressed (dark shading), after which the network maintains both memories by subsequent reactivations of the two populations. When the network is in the regime of externally generated PSs, two-item WM results in the same oscillation frequency in the global activity as that of the one-item memory (30), whereas in the regime of persistent PSs, two-item memory results in higher global frequency. In more realistic implementations of the model, an increase in frequency may be less pronounced also in the persistent regime (fig. S2).

Consequently, we propose that WM can be maintained by short-term synaptic facilitation. Accumulation of residual calcium in the presynaptic terminals could carry the information about the recalled memory in a working form, reducing the need for metabolically costly action potentials. The memories are transformed into spiking activity, either as a result of global reactivating input to the network or by virtue of the intrinsic network dynamics. Not all encountered stimuli enter WM, and we thus expect the basal modality of the network to be the transient one. The decision to allow items into WM is mediated by attention, which we suggest is represented by the global excitatory input, either in tonic or oscillating mode. The performance of human observers on memory tasks is positively correlated with the level of neural activity during the presentation of the items (34).

The model predicts that residual calcium at the presynaptic terminals should be tonically enhanced during WM, even when there is no noticeable increase in the firing rate. This prediction is in contrast to the model of (7) where WM is mediated by propagating calcium wave-

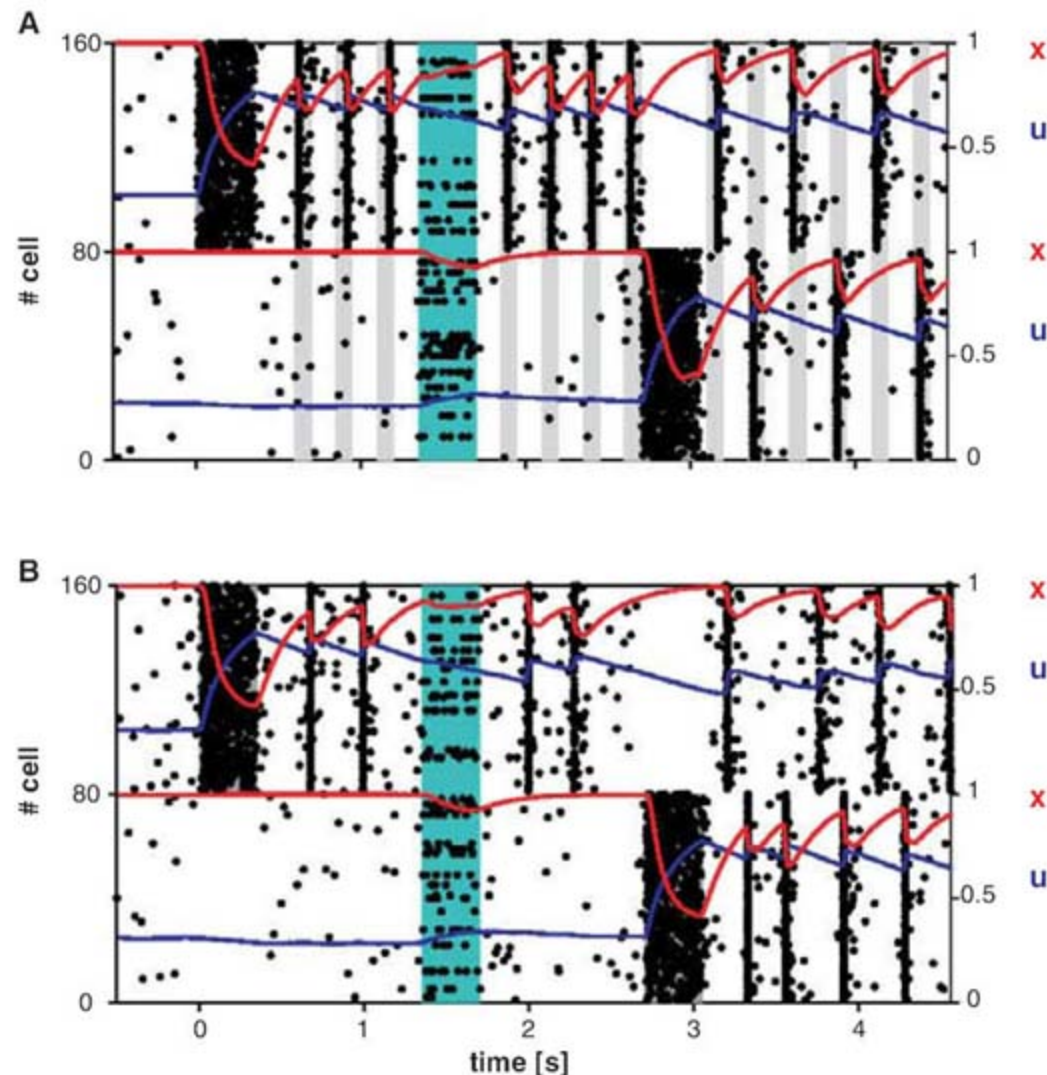


Fig. 3. Robustness to noise and two-item memory. The first item is loaded into memory at $t = 0$ (dark shading). The second item is loaded into memory at $t = 2.7$ s. Teal shading indicates a random nonspecific input to 15% of the excitatory neurons. (A) Periodic sequence of nonspecific external inputs is used to refresh the memory (gray shading). (B) Persistent PSs. Dots, rasters of 10% of the first (0 to 79) and second (80 to 159) populations' neurons; red and blue curves, same as in Fig. 2.

Biological Sample Concentration

The miVac Quattro is a versatile, high capacity biological sample concentrator that rapidly removes a wide range of solvents, from volatile organic solvents to water. It is suitable for use in applications including oligonucleotide synthesis, RNA and DNA preparation, peptide preparation, sequencing, molecular biology, and absorption, distribution, metabolism, excretion, toxicity (ADMET) research. A high displacement pump removes water quickly and quietly. Solvent removal can be enhanced by the addition of a miVac SpeedTrap refrigerated condenser. The four-swing position rotor can simultaneously accommodate up to 20 shallow-well microplates or eight deep-well plates. A wide variety of rotors is available to accommodate formats including tubes, microplates, and vials. The instrument offers digital control of concentrator temperature and short concentration times, reducing the risk of heat damage to samples.



Genevac

For information +44 1473 240000

www.genevac.com

Instrument-Free DNA Detection

The BioHelix Express Strip Cassette is a novel, disposable DNA detection device. It is designed for instrument-free, cross-contamination-proof detection of amplicons derived from polymerase chain reaction, helicase-dependent amplification, and other nucleic acid amplification reactions. The amplicons are dual-labeled by biotin and fluorescein isothiocyanate during amplification and detected on a test strip embedded in the device. The method is 10 times more sensitive than gel-based detection methods, with results viewable as colored lines in 10 minutes.

BioHelix

For information 866-800-5458

www.biohelix.com

Protein-DNA Binding Assay

The Protein-DNA Binding Assay provides a safe, fast, and sensitive alternative to traditional electromobility shift assays for detection and quantitative characterization of protein-DNA interactions. This assay eliminates the need for radioactive labeling and gel electrophoresis because it features chemiluminescent ProLabel detection and can be performed in a 96-well plate. A small (about 6 kDa) ProLabel tag is fused to the protein of interest to generate a fusion protein that is capable of producing a strong chemiluminescent signal via an enzyme complementation assay. The ProLabel tag allows direct detection of specific binding between the protein of interest and a biotinylated double-stranded-DNA oligonucleotide. The biotin moiety on the oligonucleotide permits the resulting protein-DNA complex to be captured on the 96-well plate, which is coated with streptavidin.

Clontech Laboratories

For information 800-662-2566

www.clontech.com

Bioinformatics Software

ProteinScape 2 software is a next-generation bioinformatics platform designed for addressing needs in biomarker profiling, quantification, and validation. A comprehensive solution for qualitative and quantitative liquid chromatography/mass spectrometry (LC/MS) protein analysis, ProteinScape 2 supports all current label chemistries, including multiplexed labels, as well as label-free quantification. Interactive validation of protein quantification based on raw LC/MS data is simple and straightforward. It streamlines the discovery process through decoy autovalidation algorithms and the ProteinExtractor algorithm, which produces nonredundant protein result lists across entire proteomics projects. The software has a number of dedicated data viewers that permit the evaluation and validation on each level of proteomics experiments, such as the LC/MS survey viewer, the gel viewer, and sequence MS/MS

spectra. All these views are linked and permit simple browsing through scientists' proteomics data, supported by extensive queries.

Bruker Daltonics

For information 978-667-9580

www.bdal.com

Orange Fluorescent Label

The Qdot 625 nanocrystal fluorescent label can be seen as orange and fills an important spectral gap for dense multicolor experiments. In addition, the Qdot 625 glows more brightly than longer wavelength Qdots. Current products using it include the Qdot 625 streptavidin conjugate, the Qdot 625 goat F(ab')₂ antimouse IgG conjugate, and the Qdot 625 F(ab')₂ antirabbit IgG conjugate. When excited by a single light source, Qdot nanocrystals emit in a range of brilliant, articulate colors with long-term photostability.

Invitrogen

For information 800-955-6288

www.invitrogen.com/qdots

Biodiesel Gas Chromatography

The ThermoScientific Biodiesel Gas Chromatography (GC) Productivity Solutions provide a high level of integration and productivity for the three most common GC-based methods: EN 14103 for Fatty Acid Methyl Ester characterization and linolenic acid methyl ester determination, EN 14105/ASTM D6584 for free and total glycerine content, and EN 14110 for methanol content. Delivering complete automatic characterization of pure biodiesel, the systems offer full compliance with these official standards, which aim to certify the quality of pure biodiesel for any commercial trading. The interest in biodiesel as a clean-burning alternative fuel produced from renewable sources has increased tremendously over the past few years because of its environmental advantages over conventional petrodiesel.

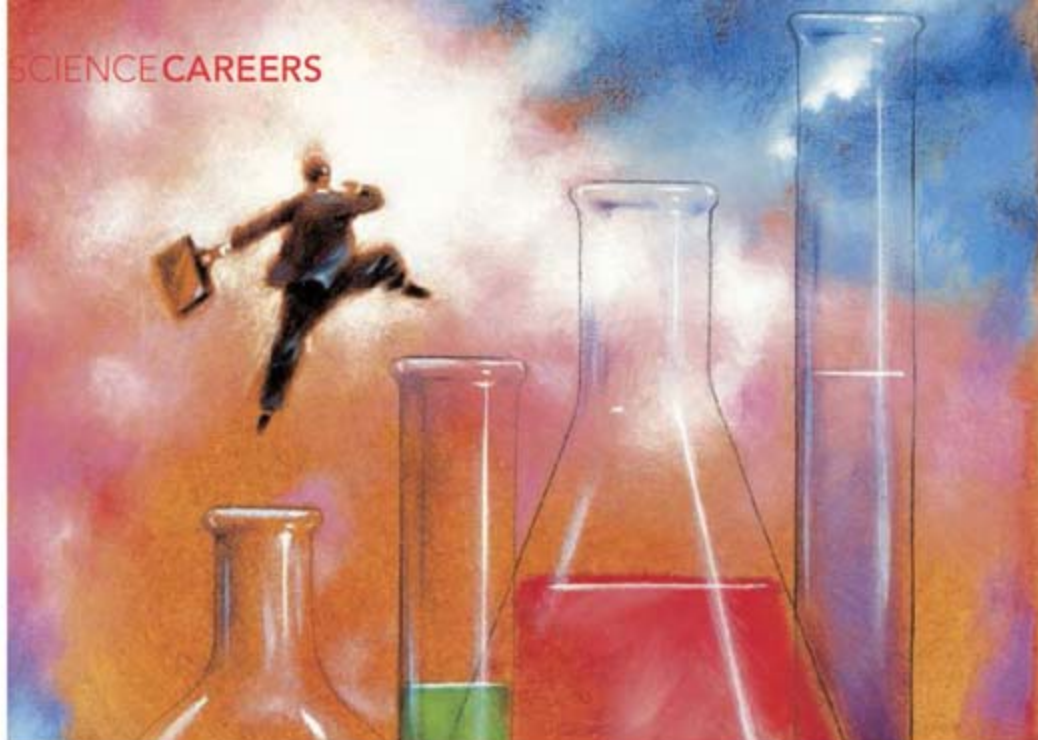
Thermo Fisher

For information 408-965-6022

www.thermofisher.com

Electronically submit your new product description or product literature information! Go to www.sciencemag.org/products/newproducts.dtl for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.



Playing Well With Industry

It's noon, and strains of Pachelbel's Canon are floating down the staircase of a Pompeu Fabra University building in Barcelona, Spain. Most visitors heading up to Xavier Serra's Music Technology Group would be surprised to find that the music comes not from a violin but from a computer. Serra feeds the computer a musical score, and the machine generates the appropriate sounds; the goal is for the music to sound like a human-operated violin. It's one of several research projects in Serra's lab that are funded by musical-instrument manufacturer Yamaha.

Serra's collaboration with Yamaha is a consequence of a natural alignment of interests; developments in music technology, he says, require this approach. He admits, however, that the money comes with some strings. Industry's primary goals—to make new products and generate money—often align well with the interests of academic researchers. But they rarely align perfectly. So, although industry is an important source of funding for academic research, industry-funded scientists have to be careful to “manage this balance between academic and industrial needs,” as Serra puts it.

The lure of industry money

In a U.S. survey conducted in 1995 by Eric Campbell, a health policy researcher at Massachusetts General Hospital and Harvard Medical School in Boston, and his colleagues, more than a quarter of life-science faculty

members reported receiving support from industry through grant agreements and research contracts. At a time of tight government budgets, the private sector's deep pockets are a provocative lure. “We are now being encouraged to do this because government funding is slowing down and the university is realizing that there must be other sources of funding to keep the research enterprise alive,” says Arijit Bose, a chemical engineer and department chair at the University of Rhode Island in Kingston, who has funding from Honda.

Industry funding fills niches that the government can't or won't fill. “Very few of the major drugs that exist today would exist if it wasn't for relationships between companies and researchers. It helps bring the results of science into the market,” says Campbell.

Ethics experts, scientists, advocacy organizations, and the public often worry about the influence of the profit motive on research. For

decades, the tobacco industry has been criticized for funding biased research and using the credentials of high-profile scientists to boost the companies' public image (*Science*, 7 January 2005, p. 36). And in January, the Center for Science in the Public Interest, a nonprofit advocacy organization in Washington, D.C., charged that U.S. universities are giving oil and gas companies too much control over the choice of on-campus research projects, first rights to intellectual property (IP), and the power to review and delay publication. The risk is especially acute in biomedical research. “People worry that relationships with industry may induce investigators into doing things such as inappropriately recruiting people in [clinical] trials and in the same way manipulating data and results,” Campbell says. Industry-funded researchers “are less likely to have results disfavoring a company.”

Industry links can also foster secrecy. “Those who are involved in academy-industry relationships are more likely to withhold data and not share data,” says Melissa Anderson, an ethics researcher at the University of Minnesota, Minneapolis. Academic scientists may agree to restrictive conditions when initiating industry relationships, and they may feel pressure later to sweep negative results under the rug. “If drugs are proved not to be effective or to have some bad effects, [there may be] pressure not to publish these findings at all,” Anderson says.

But “some of these negative effects”—bias, secrecy, and other competing interests—“are no worse than what scientists could have without industry,” Anderson says. “Every scientist must abide by the rules and policies of his or her own institution,” whether or not there is a company involved.

Industry gifts: A free lunch?

When Magdalena Balazinska, a database and information-management researcher at the University of Washington, Seattle, received a 2-year graduate fellowship, and then a New Faculty Fellowship worth \$200,000, from Microsoft Research, the money came with no IP agreements, no restrictions on her right to publish, and no research mandates. Microsoft never tried to “control what we do nor [did they] expect us to do something specific,” Balazinska says. Her experience isn't unusual: Industry gifts of reagents and equipment, fellowships, and grants that come with few strings attached are relatively common in the research world.



Other people's money. Xavier Serra (left) gets research funding from Yamaha; Magdalena Balazinska has received fellowships and grants from Microsoft and Cisco Systems.

For many researchers, however, industry money comes with certain obligations, regardless of whether they are written into formal contracts. In other 1995 results, Campbell and his colleagues found that almost half of life-sciences faculty members in the United States received gifts from industry, including biomaterials, equipment, discretionary funds, and student support. More than 40% of those receiving discretionary funds felt that they should use them only for purposes agreed to by the company, and a quarter felt an obligation to test company products. A fifth expected the company to review articles before publication, and another 14% expected the company to claim ownership of patentable results.

Whether or not there is pressure, it is in researchers' interests to please the company. Even though "there is no contract, no agreement," says Frank Dellaert, a robotics and computer-vision researcher at the Georgia Institute of Technology in Atlanta, who has received two grants of about \$40,000 from Microsoft, "it is in my own self-interest to use the money [as proposed] to build that relationship up."

Even when there are no formal contracts, most companies check on progress regularly, and some also get involved in the research.

But "I feel the same pressure with NSF"—the U.S. National Science Foundation—"because they also expect progress and [us] to work on what they fund us for," says Balazinska, who now also has corporate grants worth \$70,000 from Microsoft and Cisco Systems.

Strings

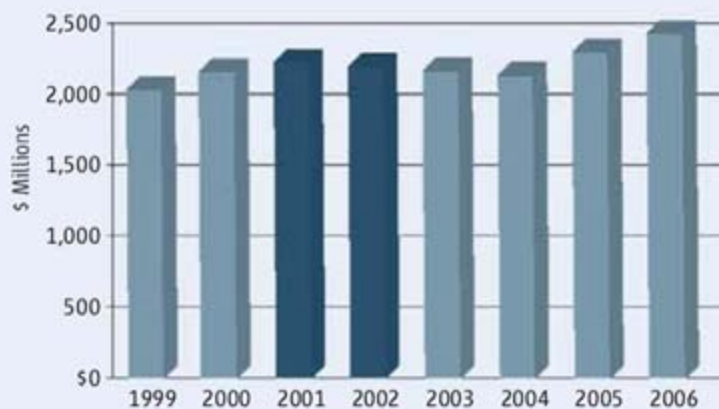
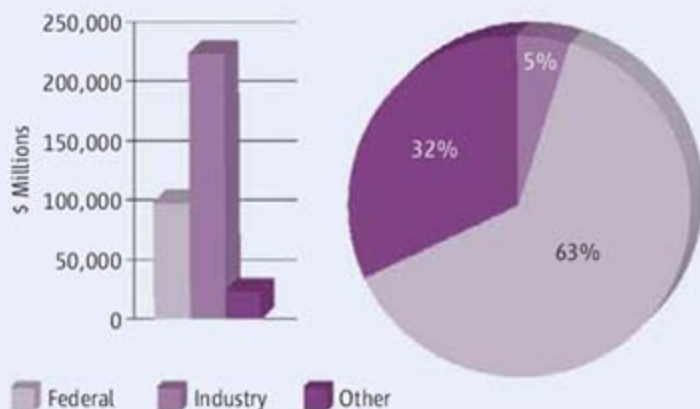
A potentially more hazardous way for academics to get corporate money is to work together with companies on projects of mutual interest, like Serra's violin synthesizer. In such collaborative agreements, academics almost always have to trade away some freedom. "Contract work is usually more specific than grant-supported research in terms of deliverables," Anderson says. "There is no question that being involved in these [collaborative] relationships is going to affect the autonomy of setting up the direction of the research. Industry can't fund people to do something that's [just] fun and interesting to them. That's one of the tradeoffs."

It's not just deciding what research to do, however. In these collaborative projects, both academia and industry contribute knowledge, effort, and resources. Both have much to gain, and to lose, so the terms and conditions of the relationship must be negotiated with care. This may mean signing away some freedom

in using the data and reporting results.

Universities, companies, governments, and professional associations all have been scratching their heads to figure out what compromises are okay. Most experts and researchers agree that accepting industry money is fine as long as academics don't compromise the academic missions of their universities. "The tipping point is when scientists' commitments to industry-funded research compromise their ability to do what their university appointments require: promote the public good through training students and pursuing research on important issues," says Anderson. "Such compromises happen when proprietary concerns get in the way of either of these responsibilities, as in delaying a student's progress toward a degree or deflecting a research program from important, broad questions to questions of little concern outside a company's specific development program."

One of the most contentious issues is ownership of IP. According to the U.S. Business-Higher Education Forum, an organization of Fortune 500 CEOs, college and university presidents, and foundation leaders, an appropriate arrangement is for universities to keep IP ownership while allowing companies to commercialize innovations. A common solution is to grant the



Business Financing for Your Research

Scientists at academic institutions usually turn to government and foundations for money to do their research, and for good reason: These sources provide the bulk of the funding for academic research. But the fluctuating federal research budgets of recent years, in Europe and Japan as well as the United States, have given researchers reasons to go looking for alternative funding sources.

Business spends more money on research and development than the public and academic sectors combined (bar chart, above left). But only a small proportion of that investment reaches academic researchers. In 2006, according to figures compiled by the U.S. National Science Foundation, industry spent \$2.4 billion at U.S. academic insti-

tutions. That's just 5% of total funding for academic research (pie chart, above), but it's still a nice chunk of change.

Speaking of change: That number hasn't changed much over time, in either direction. Businesses have roughly maintained their levels of academic research support (in current dollars) since 1999, declining only slightly in the years of slow economic growth that followed the 2001 recession (bar chart, above right). Level funding (or a small decline) doesn't sound impressive in an era when researchers aspire to doubling research budgets, but steady private-sector funding can provide a small cushion given the ups and downs in government support for science.

—ALAN KOTOK

Alan Kotok is managing editor of *Science Careers*.

company exclusive or nonexclusive licensing rights or the right of first refusal.

Another thorny issue is data ownership and disclosure. As the Federation of American Societies for Experimental Biology (FASEB) has stated, any terms preventing publication of research results or asking for ghost authorship are unacceptable.

Companies are routinely given the option to review publications for confidential information and patentable data. In the United Kingdom, the government has provided a set of legal contracts—the Lambert Agreements—for researchers to draw on as models for reasonable compromises with industry. Under the terms of these agreements, companies are given up to 30 days to review manuscripts. Companies may also ask academics to keep new data secret until a patent application is filed. The Lambert Agreements give companies up to 12 months for this, but the U.S. Business–Higher Education Forum recommends limiting this delay to 90 days.

Model agreements and guidelines can be very helpful, but things can go wrong. That's

one reason most universities now have business-development and technology-transfer offices; it's a good idea to get them involved in any deal as early as possible. Indeed, many institutions require professors to disclose links with industry so that they can oversee those relationships and help researchers avoid sticky situations. "You have to go through university lawyers who see that all the university's and investigators' interests are protected," as well as the integrity of science, says Elizabeth Heitman, a biomedical ethicist at Vanderbilt University in Nashville, Tennessee. According to a white paper from the National Academies' Government-University-Industry Research Roundtable, in the United States, such negotiations on average take 70 days and may last as long as several years.

If you do manage to navigate your way to a deal and produce publishable results, you have one more important obligation: Most journals require authors to disclose conflicts of interest. Campbell also recommends that researchers disclose ties with industry whenever they present their results.

Recognition and career prospects

According to the U.S. Business–Higher Education Forum, industry-funded students may take an extra 6 months to earn their Ph.D.s compared with publicly funded peers. Yet industry involvement can be an excellent door opener. Industry grants and fellowships "are probably fine as long as the student has the opportunity to choose his or her dissertation [topic] and conduct the research independently of the company and publish their results," Campbell says. Academic supervisors must assure that the projects are Ph.D.-worthy, that graduation and publications aren't delayed unreasonably, and that students are aware of the potential pitfalls.

Experience working with industry may be especially beneficial for trainees who expect to work in industry. Postdocs seeking jobs in academia have less space to maneuver. They cannot afford to have their publication record put on hold and should make sure any delays are offset by a constant stream of papers from other, publicly funded projects. "If you are an early-career researcher and too much into patenting rather than publishing, your

Finding Industry Funding

Most collaborations between companies and academic researchers are initiated by industry scientists looking for specific technologies or expertise, notes Anthony Boccanfuso, executive director of the Washington, D.C.-based University-Industry Demonstration Partnership (UIDP), a consortium aimed at improving collaboration between universities and industry. But that doesn't mean university researchers have to sit and wait for the phone to ring. If you are looking for industry funding, it helps to be proactive.

Exploit ready-made opportunities

In many cases, researchers interested in private-sector funding need look no further than their own campuses. Many research institutions already host industry-funded programs. At their most ambitious, these programs are massive, multicenter research consortia that recruit dozens or even hundreds of industry partners. Albany NanoTech, a multibillion-dollar research complex affiliated with the College of Nanoscale Science and Engineering at the University at Albany in New York state, involves more than 250 corporations, many of which provide major funding for university faculty members.

Many universities also administer industry-funded grant programs on a smaller scale. ConocoPhillips, for example, recently announced that it will give \$22.5 million over 8 years to the multidisciplinary biorenewable fuels research program at Iowa State University in Ames. The grant will fund about 10 faculty members in its first year.

Tune in to industry needs

Establishing new university-industry alliances requires some legwork. If you have a good sense of which companies might profit from your work, search those companies' Web sites for faculty fellowships. Also look for competitive requests for proposals (RFPs), a mechanism a small but grow-



Take the lead. Anthony Boccanfuso (left) urges academic scientists to make their case to industry, as Karen Wooley of Washington University in St. Louis has done.

ing number of companies use to provide grants to academic researchers.

For the past 2 years, a response to an RFP has netted two unrestricted grants of about \$40,000 for computer scientist Frank Dellaert, who studies robotics and computer vision at the Georgia Institute of Technology in Atlanta, to develop new online three-dimensional mapping technologies for Microsoft's Virtual Earth. Dellaert says the RFP application process is far less cumbersome than some federal grant applications, which require technical proposals 15 to 60 pages long. "With Microsoft, you write one page of text; there is no budget, just a back-of-the-envelope calculation. It's extremely painless."

In addition to developing formal RFP procedures, many large corporations—DuPont and Pfizer, for example—have adopted an R&D model known as open innovation, forming corporate "technology connectors" to fund external scientists and entrepreneurs to work on tightly defined problems. A consumer-products company, for example, might use this

progress toward an academic career can be compromised," Anderson says.

Industry collaborations may look like gold stars on the CVs of young professors. "Usually it's only the scientists who are doing really well in academic circles" that industry seeks out for collaborations, Anderson says. Yet, industry grants are perceived as "not peer-reviewed and not as prestigious" as public funding, Campbell says, so they may not be given as much weight in the tenure process. Furthermore, the strings attached—obstacles to publishing and presenting your work and limitations on your ability to interact with other researchers—can hold back your progress toward tenure and promotion.

This is a price Serra is willing to pay—and has. "I am not a full professor," he says. If he had worked less with industry and more with traditional funding sources, "I would have been a long time ago." But the opportunity to work with a company like Yamaha at the interface of science, engineering, and music justifies the sacrifice. "I lose certain things; I gain others," he says.

—ELISABETH PAIN

Elisabeth Pain is a contributing editor for *Science Careers*.



Additional Resources

A roundup of the issues in the Business–Higher Education Forum's *Working Together, Creating Knowledge* report

www.acenet.edu/bookstore/pdf/working-together.pdf

Practical guidance from the U.S. University-Industry Demonstration Partnership

www.uidp.org/

The European Commission's IPR Helpdesk

www.ipr-helpdesk.org/controlador/principal?seccion=principal&len=en

Responsible Partnering in Europe

www.responsible-partnering.org/

The U.K. Lambert Agreements

www.innovation.gov.uk/lambertagreements/

The FASEB tool kit for managing financial relationships with industry

opa.faseb.org/pages/advocacy/coi/About.html

The Committee on Publication Ethics guidelines on good publication and the code of conduct

www.publicationethics.org.uk/guidelines

mechanism to request proposals for a polymer that adheres to polyester surfaces. Open innovation is largely focused on technology transfer or intellectual-property licensing, but opportunities for funding academic research also exist.

Some companies also work through intermediaries, such as the companies NineSigma and InnoCentive, which broadcast their clients' R&D requests to scientists, engineers, and entrepreneurs around the world who have opted into the companies' proprietary "solution provider" networks.

Build relationships

Too often, academic researchers seeking corporate support for their work try to convince industry scientists that what they do is great science, notes Michael Amiridis, a chemical engineer and dean of the college of engineering and computing at the University of South Carolina, Columbia. Instead, he says, "try to understand what the problem is that industry is trying to solve and show what you can bring to the table. It takes cultivation, and it takes time."

"A lot of people think, 'Oh, companies have lots of money; let's get some of that money,'" says chemist Karen Wooley of Washington University in St. Louis, Missouri, who studies nanometer-scale polymer particles and has received funding from a number of industry sources. "That's a naive point of view. We can't expect handouts or open-ended gifts. It has to be a friendship based on mutual respect, and there has to be an indication that the company is going to get something out of it."

Making personal connections is "far and away the best way to get a deal done," says David Rosen, executive director in worldwide business development at Pfizer. "Getting a champion inside the company to want your technology breaks through a lot of barriers. If somebody inside the company is pulling while you're pushing, stuff tends to get done a lot faster and a lot more effectively."

The trick is knowing where to start. Many academic researchers feel disconnected from the private sector, and targeting companies "cold" is "a

tortured path because companies' internal structures aren't usually transparent to the public," notes Boccanfuso.

But finding industry scientists with compatible interests isn't as hard or as mysterious as it might seem. Industry researchers graduate from the same doctoral programs that generate academic researchers, belong to the same scientific societies, attend the same conferences, publish in the same journals, and register with the same patent office. That means that with some detective work, you can figure out who they are.

At professional conferences, take advantage of opportunities to meet industry scientists. "I see the tendency of academics to cluster among themselves at these meetings," says Amiridis. "I don't do this. I go out and look at the nametags. Talk to people. You need to sell yourself." While you're chatting, nurture these fledgling relationships by inviting your new industry friend to give a seminar in your department. Return the honor by offering to give a seminar to the company's scientists.

Make yourself as visible as possible. Set up a profile on the professional networking site LinkedIn and indicate in your contact settings that you're interested in consulting offers, job inquiries, and expertise requests. Make full use of university databases that detail researchers' interests, suggests cognitive psychologist Dennis Proffitt of the University of Virginia, Charlottesville, who has received funding from several industry sources.

University research administrators can also be rainmakers, says Don Gerhart, associate vice president for research and innovation at the University of Oregon, Eugene. Many research administrators, he notes, monitor opportunities for research partnerships with the private sector and are happy to help guide faculty members toward industry collaborations.

When all else fails, pick up the phone and make cold calls. "Be bold and invite yourself in, even if it's going to cost you some money," says Amiridis. "Maybe three out of four times it won't lead to anything, but the one time that it does, it can easily pay for the other times."

—SIRI CARPENTER

Siri Carpenter is a freelance science writer in Madison, Wisconsin.

YIELD



SPEED

**WE THINK YOU
SHOULD BE ABLE
TO OPTIMIZE
ON THE FLY**

Convenience Is What Happens When You Rethink PCR

Change the way you think about PCR with Bio-Rad's new family of thermal cyclers.

Wouldn't you rather optimize your reactions in minutes and not days? With Bio-Rad's new 1000-series thermal cyclers, optimizing on the fly is just the beginning.

- Easily interchangeable reaction modules meet any experimental or throughput need
- Reduced-mass sample blocks increase ramp rates and reduce run times
- Thermal gradient lets you incubate each row at a different temperature for fast protocol optimization

When you rethink PCR, you realize how easy it should be.

For more information, visit us on the Web at www.bio-rad.com/pcr



Purchase of this instrument conveys a limited non-transferable immunity from suit for the purchaser's own internal research and development and for use in applied fields other than Human In Vitro Diagnostics under one or more of U.S. Patents Nos. 5,656,493, 5,333,675, 5,475,610 (claims 1, 44, 158, 160-163 and 167 only), and 6,703,236 (claims 1-7 only), or corresponding claims in their non-U.S. counterparts, owned by Applied Biosystems. No right is conveyed expressly, by implication or by estoppel under any other patent claim, such as claims to apparatus, reagents, kits, or methods such as 5' nuclease methods. Further information on purchasing licenses may be obtained by contacting the Director of Licensing, Applied Biosystems, 850 Lincoln Centre Drive, Foster City, California 94404, USA.